Steps and resources for submitting data using WQX

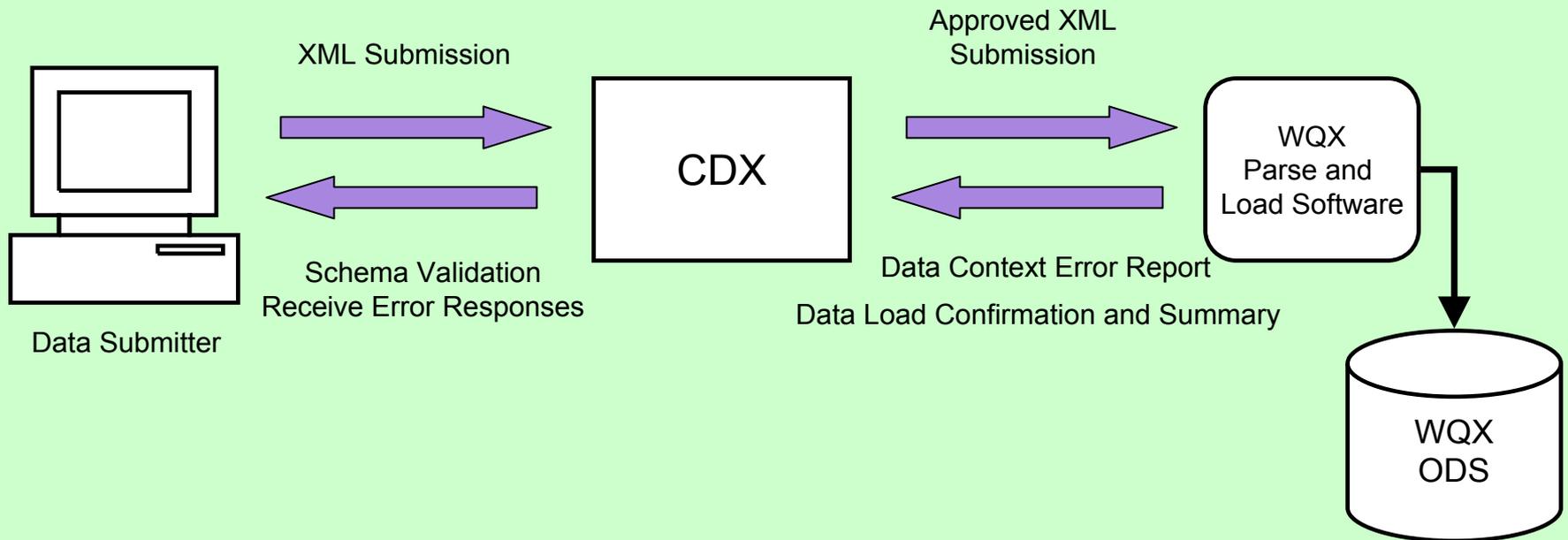R9 STORET/WQX Workshop for Tribes

November 19, 2008

Submitting data using WQX

- Water Quality Exchange (WQX) uses the Exchange Network to flow data from an tribe's own database to the existing STORET Warehouse

- Creating XML is an important piece

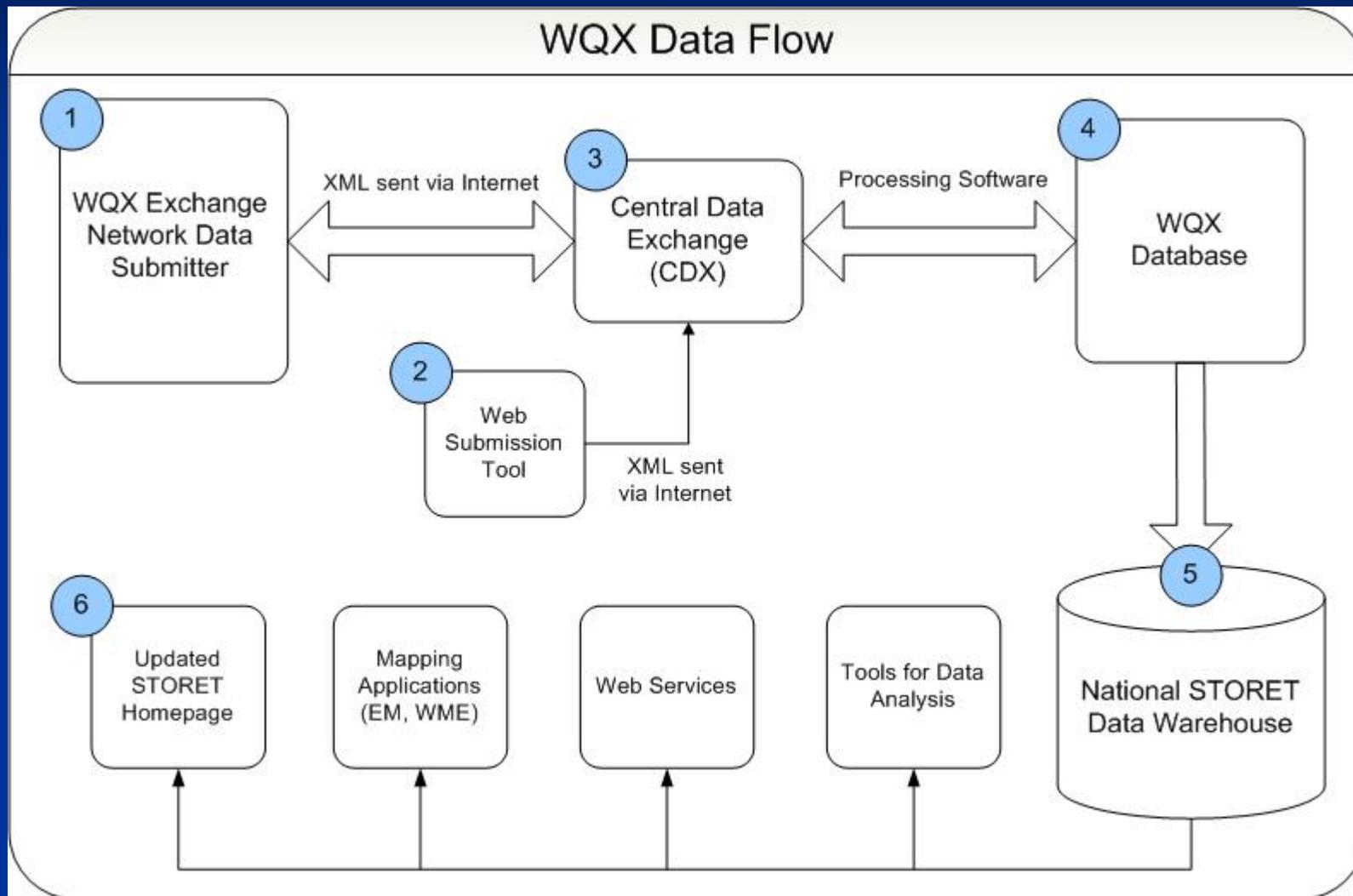- A database data management system is critical though not necessary

What makes up the WQX flow?

- **Data Partners**
- The WQX Schema
- The WQX Flow Configuration
  - WQX Processing Software
- The STORET Warehouse
- The National Environmental Information Exchange Network
- The Central Data Exchange

How WQX Works

XML Submission

Approved XML Submission

CDX

WQX Parse and Load Software

Schema Validation
Receive Error Responses

Data Submitter

Data Context Error Report

Data Load Confirmation and Summary

WQX ODS

WQX in the big picture

WQX Data Flow

1 — WQX Exchange Network Data Submitter

XML sent via Internet

3 — Central Data Exchange (CDX)

Processing Software

4 — WQX Database

2 — Web Submission Tool

XML sent via Internet

6 — Updated STORET Homepage

Mapping Applications (EM, WME)

Web Services

Tools for Data Analysis

5 — National STORET Data Warehouse

Moving data using WQX

1. Review the resources that are available
2. Getting involved with the Exchange Network
3. Adopting the WQX schema as a standard way of sharing data
   1. WQX Data Elements and Data rules
   2. WQX Domain Values
4. Understanding how your data relates to the WQX structure and the XML file type
5. Submitting a file
6. Understanding errors
7. Interacting with WQX

Step 1. Resources Available

- The STORET/WQX team has several resources available to help you:
  - Available at the STORET web site: www.epa.gov/storet
    - Visit the WQX section: http://www.epa.gov/storet/wqx.html
    - Diagrams and presentations about how WQX works
    - Presentations from others on how they've implemented WQX (including tribes)
    - Lists of allowable values for WQX
  - Read the Flow Configuration Document
    - See example XML Files
   - Examine the Data Exchange Template

Step 1. Remember this

All of this information is available at:
www.epa.gov/storet

Step 2. Getting involved in the Exchange Network

- Ask yourself the following questions:

  1. Does my organization have a node?
  2. Have we submitted data for another flow?
  3. Did we use a node? Node client? Or CDX Web?
  4. Who is my Node Administrator? Do we have one?

- If the answer to these questions is 'No' or 'I don't know' visit: www.exchangenetwork.net

What is the Exchange Network?
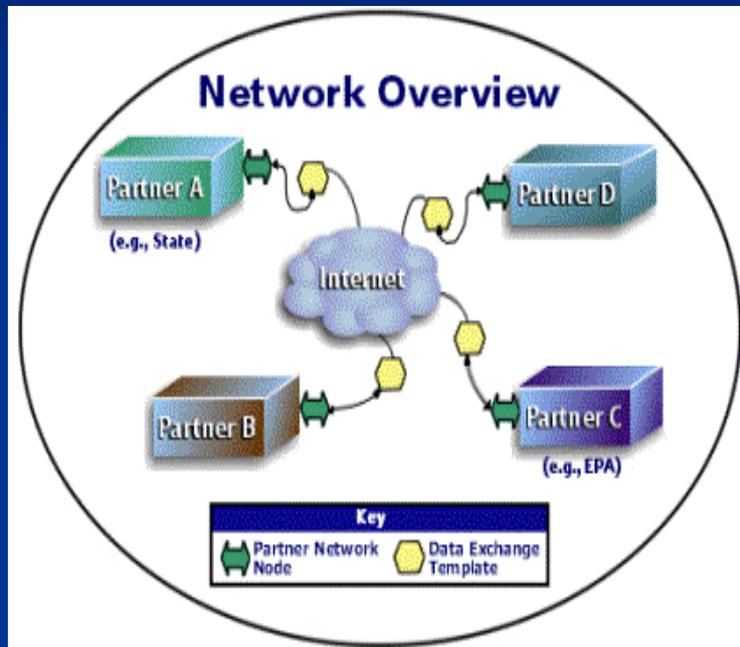
Principles:

1. Each partner manages their own data, and is the steward for data they originate.

2. The Network "doesn't care" how a partner manages and stores information.

3. Partners who need to store copies of data locally are responsible for the design, format, and maintenance of the copy they create.

4. Network partners agree to use adopted technology and data standards.

A common approach to sharing data that lets you use your partners' data as if it is a part of your own systems.

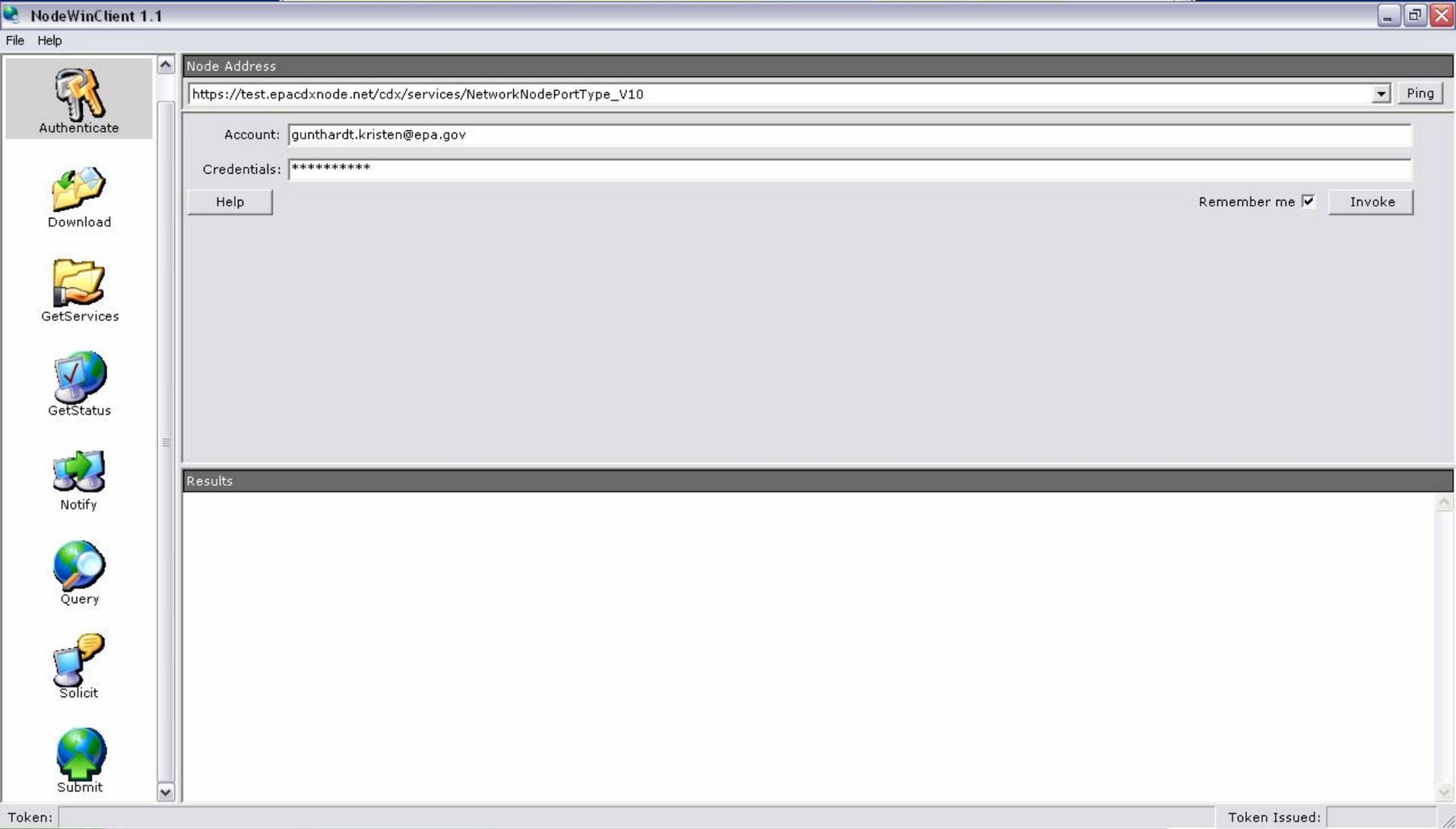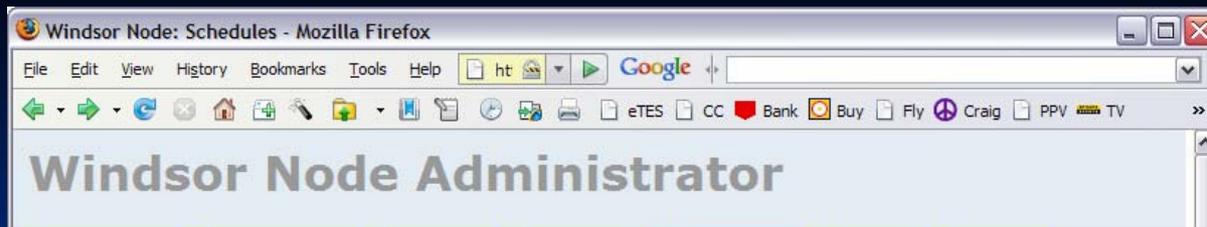# Nodes and Node Clients - the communication devices of the EN

Node

Node Client

- Makes Requests of other Nodes using XML
- Responds to Requests
- Comes with a closet

- Makes Requests of other Nodes using XML
- Responds to spilt liquid
- Comes with a monitor

Step 3 – Adopting the WQX Schema

- Available at the Exchange Network Site: http://www.exchangenetwork.net/exchanges/water/wqx.htm

- Schema is a computer definition of the WQX data elements

- Data Exchange Template (DET) is a spreadsheet that lists out all of the WQX data elements and their business rules

- Domain Values are posted on the STORET website: http://www.epa.gov/storet/wqx_downloads.html

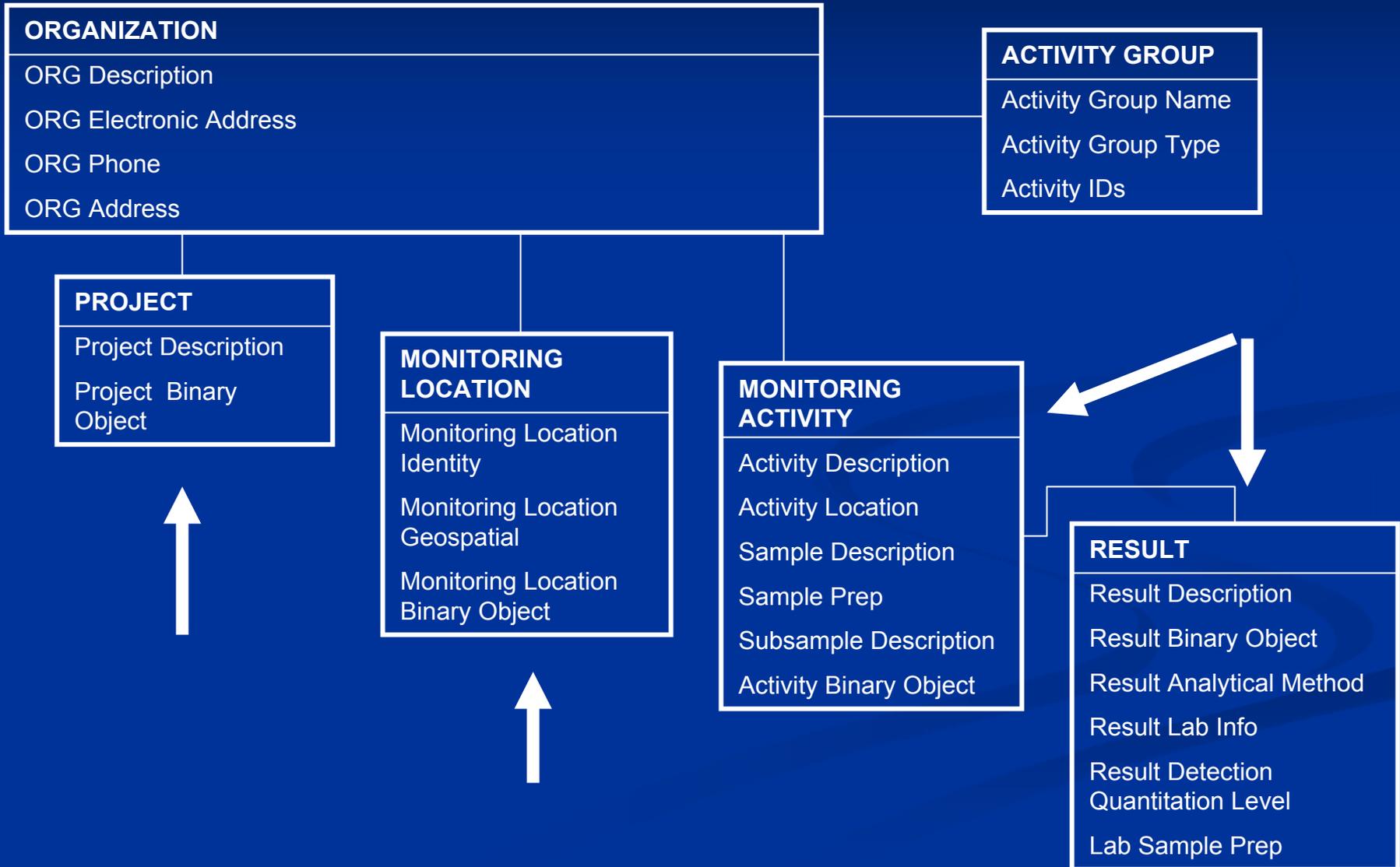- Domain Values are also available via "WQX.GetDomainValueByElementName"

WQX Schema

- A definition of the structure and format that data must have to be submitted through WQX.

- Common set of data elements that everyone uses when submitting data

- Primary components: Organization, Project, Monitoring Location, Monitoring Activity, Samples, and Results

- Database independent – xml can be created out of any type of database

How EPA Describes Monitoring Data

Three categories of data:

- **Projects** (Why data were collected)

  - Brief summary of monitoring plan

- **Monitoring Locations** (Where data were collected)

  - Describe where monitoring takes place

- **Results** (When, How and What were collected)

  - Measurements of what were monitored

WQX is designed around these concepts

ORGANIZATION

ORG Description

ORG Electronic Address

ORG Phone

ORG Address

ACTIVITY GROUP

Activity Group Name

Activity Group Type

Activity IDs

PROJECT

Project Description

Project Binary Object

MONITORING LOCATION

Monitoring Location Identity

Monitoring Location Geospatial

Monitoring Location Binary Object

MONITORING ACTIVITY

Activity Description

Activity Location

Sample Description

Sample Prep

Subsample Description

Activity Binary Object

RESULT

Result Description

Result Binary Object

Result Analytical Method

Result Lab Info

Result Detection Quantitation Level

Lab Sample Prep

WQX schema v1.0

- The physical conditions in the environment at the time of a site visit.

- The chemical and bacteriological make-up of the water sampled.

- Chemical analyses of fish tissue collected.

WQX schema v2.0 – What's new?

- Biological Taxon Abundance data, including population census, frequency class, group summaries, and individual results

- Reference site information

- Toxicity data

- Habitat Assessment scores and their related metric scores

- Biological Index scores and their related metric scores

Step 4. Understanding the WQX Structure

- WQX files should be organized as follows:
  - Header (with processing instructions)
    - Organization
      - Organizational Information
      - Projects
      - Monitoring Locations
      - Activities
        - Linkage to Projects
        - Linkage to Monitoring Locations
        - Results
      - Activity Groups

Step 4 (cont). Understanding XML

- XML can be intimidating at first
- There are many ways to create XML, some things to do:
  1. The sample files and Data Exchange Template are key to understanding how to create a WQX XML file.
  2. Find a programmer who can spend 20-40 hours on designing an XML file
  3. Have some test data that are representative of the data you'll be submitting
  4. Test the procedures that the programmer developed (plan for an additional 20-40 hrs testing)
  5. Start small, it takes time to create an XML file, don't waste development time waiting for a computer to process a file.

***You should only have to do this once, after that, you're just running the procedure.

XML in a nutshell

```
- <Project>
    <ProjectIdentifier>11</ProjectIdentifier>
    <ProjectName>2005 Acme River Reach</ProjectName>
    <ProjectDescriptionText>River Reach conducted in 2005</ProjectDescriptionText>
  </Project>
+ <MonitoringLocation>
+ <MonitoringLocation>
- <Activity>
  + <ActivityDescription>
  + <ActivityLocation>
  + <SampleDescription>
  - <Result>
    - <ResultDescription>
        <DataLoggerLineName>1</DataLoggerLineName>
        <ResultDetectionConditionText>Present Below Quantification
          Limit</ResultDetectionConditionText>
        <CharacteristicName>3-Methyl-1,1'-biphenyl</CharacteristicName>
        <ResultSampleFractionText>Total</ResultSampleFractionText>
      - <ResultMeasure>
          <ResultMeasureValue />
          <MeasureUnitCode />
          <MeasureQualifierCode />
        </ResultMeasure>
```

- eXtensible Markup Language
- An "XML Document" contains the Data
- An "XML Schema" defines the Structure and the Rules of the data
- The WQX schema is the computer definition of the Structure and Rules for WQX data

Step 5. Submitting a File

- Read the WQX Flow Configuration Document
- Set up a NAAS account and an Org ID, and successfully created an XML file for submission
- The original test file should be small, but representative of the data you'll be submitting
- Submit to CDX test first
- Once you get a successful submission, keep track of the transaction ID
- Send an email to the CDX help desk, requesting access to WQX production. Provide them the transaction ID of the successful submission.
- Go ahead an begin submitting to production

Getting a NAAS Account

- The NAAS or Network Authentication Authorization Service is a user account that you can use to access the Exchange Network. They can be specific for an individual node or for a particular user

- NAAS accounts are assigned by Node Administrators. If you don't have a Node Administrator, contact the CDX help desk at: 888-890-1995 Or epacdx@csc.com

Getting an Org ID

- An Organization ID (Org ID) is a unique identifier assigned to your organization (i.e. YUROK for the Yurok Tribe).

- Org IDs are assigned by EPA upon the request of the user.

- You must have an Org ID to be able to submit to WQX

- To request an Org ID, send an email to storet@epa.gov
  - Provide the NAAS accounts that you want to have access to your Organization

Step 6. Understanding Errors

- There are two types of errors in WQX:

    1. **Schema Errors**: the submitted file does not conform to the WQX Schema. (Report is available in the file 'Validation Results' which is an xml file).

    2. **Processing Errors**: the submitted file does not pass all of the WQX Business Rules (Report is available in the file 'ProcessingReport.zip')

Step 7. Interacting with WQX

- WQX is designed to function in a Node-to-Node environment (i.e. you manage your data, let your computer and Node keep EPA in sync).

- WQX allows Inserts, Updates, and Deletes via a batch process

- WQX has Outbound services via the Query and Solicit method data visualization and system synchronization purposes.

STORET Warehouse

- Once data has been submitted using WQX, it is placed into the STORET Warehouse at the end of every week

- All results must be marked "Final" in order to be placed in the Warehouse

- E.g. I submitted data through WQX on November 17th at 10 am. My data will be in the STORET Warehouse by November 24th.

HELP

- Last but not least, don't be afraid to ask for help:
  - Join the STORET list serve: www.epa.gov/storet/listserv.html
  - Call the STORET hot line: 1-800-424-9067
  - send an email to storet@epa.gov