LOCKHEED MARTIN
We never forget who we're working for

Information Technology Solutions - Environmental Systems Engineering
(ITS-ESE)
Contract No.: 68-W-04-005
Task Order: 08

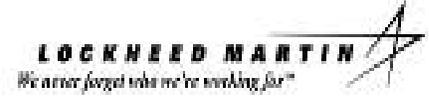**Office of Water, Water Quality Exchange (OWWQX) Pilot**

Lessons Learned Report

Version Control

Date	Author	Changes	Version
03/15/06	E. Bryant	Draft	1.0
03/16/06	K. McNeill	Edit/Review	1.0
04/03/06	E.Bryant	Edit-Final Review	1.0

Table of Contents

Introduction

Background

The U.S. Environmental Protection Agency's (EPA's) Office of Information Collection (OIC) and Office of Wetlands, Oceans and Watersheds (OWOW) is committed to implementing Central Data Exchange (CDX) services and establishing the EPA infrastructure to support an ambient water quality monitoring data exchange. The Water Quality Monitoring data exchange project is the product of a collaborative effort between OIC, Office of Water (OW), and the Environmental Council of States (ECOS). The project was identified during the development of the Environmental Sampling Analysis and Results (ESAR) data standard for water monitoring.

The project goal is to provide EPA state partners with a means of exchanging water quality monitoring data via CDX, using the Office of Water, Water Quality Exchange (OWWQX) data standard (a hybrid of the draft ESAR data standard as it stood at the time of the pilot initiation). OW, in partnership with the states, will establish water quality monitoring data exchange elements, business rules for exchanging these elements, and valid domain lists for elements not covered by an existing or proposed standard.

The Office of Water, Water Quality Exchange (OWWQX) Pilot established a data flow through which three initial pilot states (Oregon, Michigan, and Texas), and a tribal organization, the Wind River Environmental Quality Commission (WREQC). In all cases the goal was to submit Water Quality Monitoring (WQM) data to EPA via the CDX Exchange Network node.[1]

The OWWQX pilot included the workflow for both CDX Node and CDX Web submissions. Only four nodes were involved in this initial pilot, namely: Michigan, Oregon, Texas, and WREQC. The pilot was implemented in the CDX Pre-Production environment.

Exhibit 1 depicts the high level overview of the OWWQX Pilot system. The OWWQX Pilot system consists of the following elements:

- Nodes - Michigan, Oregon, Texas, and WREQC
- CDX
- OWWQX Back-end Application
- OWWQX Database

---

[1] *Note: WREQC is a joint commission of the Arapaho and Shoshone tribes.*

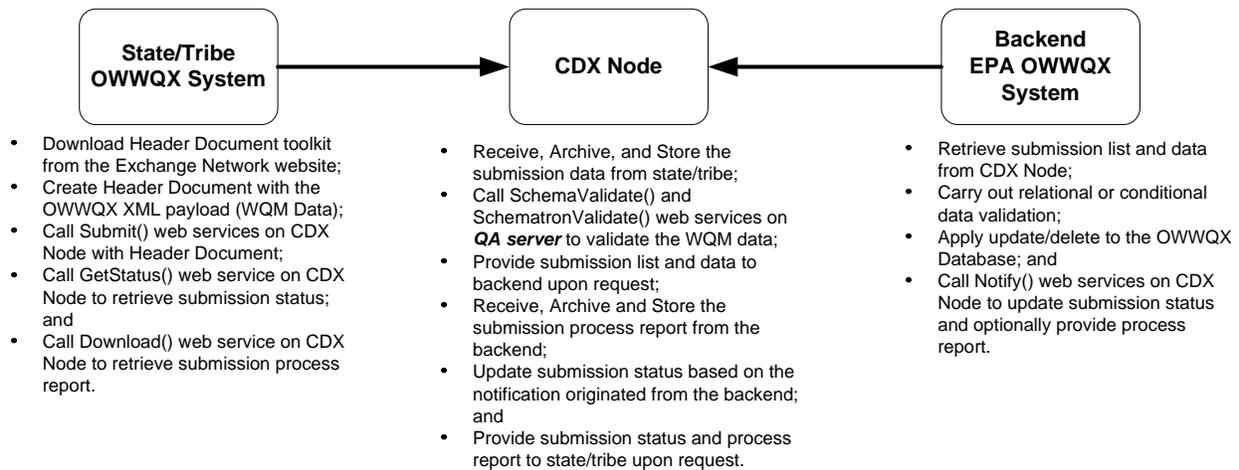| State/Tribe OWWQX System	CDX Node	Backend EPA OWWQX System
• Download Header Document toolkit from the Exchange Network website; • Create Header Document with the OWWQX XML payload (WQM Data); • Call Submit() web services on CDX Node with Header Document; • Call GetStatus() web service on CDX Node to retrieve submission status; and • Call Download() web service on CDX Node to retrieve submission process report.	• Receive, Archive, and Store the submission data from state/tribe; • Call SchemaValidate() and SchematronValidate() web services on **QA server** to validate the WQM data; • Provide submission list and data to backend upon request; • Receive, Archive and Store the submission process report from the backend; • Update submission status based on the notification originated from the backend; and • Provide submission status and process report to state/tribe upon request.	• Retrieve submission list and data from CDX Node; • Carry out relational or conditional data validation; • Apply update/delete to the OWWQX Database; and • Call Notify() web services on CDX Node to update submission status and optionally provide process report.

Exhibit 1 – OWWQX Pilot System Overview

The OWWQX pilot included submitting the following types of data:

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

- Chemical and bacteriological make-up of the water sampled.

- Optionally, chemical analyses of the tissues of any fish collected.

The pilot data flow utilized two mechanisms for exchanging water quality monitoring information; a Web-based solution for manual submissions and a Web services-based solution for automated submission utilizing Exchange Network standards. Upon EPA approval of the pilot data flow, the data flow will be deployed to CDX Production.

Document Purpose

The purpose of this document is as follows:

- Summarize the pilot project.

- Discuss key pilot aspects and lessons learned.

- Provide the OWWQX issues.

This document is intended as a historical recollection of the OWWQX pilot project, as well as a technical reference for the design of the Water Quality Monitoring production system. It is a reflection on what was done right in the pilot project development phase, what could have been done differently, and how to be more effective in the future.

Summary of Pilot

Major Deliverables and Milestones

The OWWQX pilot project kicked – off in mid April 2005 and concluded in late January 2006. The major deliverables associated with the project activities are listed in the following table.

Product	Completion Date
XML Schema and Schematron Files	3/2005
Domain Validation Lists	3/2005
Design Operational Data Store (ODS)	6/2005
Develop draft flow configuration document	7/2005
Deployment of QA Services	8/2005
Establishment of CDX and ODS test environments for OWWQX	8/2005
Develop ODS installation package	10/2005
Development of data processing software for parsing and loading of XML documents	10/2005
Integration of CDX and ODS test environments	10/2005
Mechanism for returning data load errors and transaction receipts	11/2005
Flow data between pilot participants and EPA	01/2006

Key Pilot Aspects and Issues

The OWWQX bi-weekly meetings were used to achieve consensus, promote exploration, and ensured detailed documentation of the process and the underlying data structure. The participants understood how active collaboration focused on a shared goal could yield high quality results. Action items were assigned and tracked at the bi-weekly teleconferences and issues/concerns were discussed in detail. The bi-weekly meeting minutes are presented in the Appendix of this document. This section covers the key aspects of the pilot and emerging issues related to the activities.

Schema

- Insert, Update and Delete functionality supported via two schemas (one for Update-Insert and one for Delete)
  - o An Update-Insert Submission must include a complete data set.
    - For example: OWWQX does not support a submission containing Results only. All Projects, Monitoring Locations, etc. that are referenced, must also be included in the file.
    - This is partly to support Schematron Validation which doesn't have access to the database (to validate a Project ID, for example).
- Organization
  - o A restriction was added to only allow one Organization per submission.
- ActivityIdentifier
  - o The decision was made to require that ActivityIdentifier be unique within an Organization
    - It may be inconvenient for some data providers to generate unique IDs for Activities, but that this was still probably the best option under the circumstances.
  - o Originally an Activity was uniquely identified by the following elements: OrganizationIdentifier, MonitoringLocationIdentifier, ActivityStartDate, and ActivityIdentifier
    - Having a composite key like this made it impossible (with the existing schema) to submit an 'update' of an Activity if the Monitoring Location or Activity Start Date had changed (because the previous values would be needed to find the existing Activity).

- - The schema would have to be changed in awkward ways to address this and might prevent the sharing of the same schema for Insert and Update.
    - Likewise, the auditing burden on a data provider would be higher because they would have to keep historical values for the Monitoring Location and Activity Start Date on an Activity anytime it changed in order to correctly submit the Activity Update information to OWWQX.
- Handling of sample and subsample information changed during the pilot – before, sample method, prep, preservation and transport information was required for both sample and associated subsamples. Now, this information is tied to a separate sample prep table that both sample and subsample records point to.
- AttachedBinaryObjects
    - It was determined that AttachedBinaryObjects (gif, pdf, etc) were to be submitted as separate files rather than embedded into the XML submission.
        - The schema has an element intended to hold an embedded object (i.e. BinaryObjectContent). This was left in the schema for the pilot, in case the issue was readdressed in the final implementation.
        - If the decision for the final implementation is to only allow binary objects as separate files, then the BinaryObjectContent element in the schema can be dropped.
- Required elements
    - Originally the OWWQX XML schema defined required elements using minOccurs="1". This only required that the tags be included in the XML document (allowing empty tags to pass schema validation).
    - The schema was later updated to also use minLength value="1" for required elements so that empty tags would not be valid for required elements.

Operational Data Store (ODS)

- Created with Oracle 9i
- Used local tables for domain values (no syncing with System of Registries [SoR])
    - A decision was made to delay integration with the SoR (for Domain Lists) until after the Pilot.
- Auditing Infrastructure
    - Although auditing was not a priority for the Pilot, basic auditing functionality was implemented. Currently the ODS has a TRANSACTION_LOG table that tracks basic information about each Insert, Update, and Delete that is made (while loading an XML Submission). This information proved useful in testing the application, by proving a way to confirm that the changes that were expected, while loading a file, actually occurred. It was also used (in summary form) in the Processing Report sent back to a submitter (e.g. summary counts of Results inserted, updated, or deleted).
    - Some enhancement to this might be considered in the final implementation.

Parse and Load

- Created with C# .NET 2003
- Limited data validation
    - Because of the validation that occurs at CDX, only limited data validation was placed in the Parse-and-Load software.
- Zipped vs. Standard XML files for submission
    - Originally data submissions were to be submitted as Zip files.

- XML files are much more efficiently transmitted when zipped because they compress significantly (due to a high number of repeated tags).
- Zipped files also accommodate the external binary object files as well by providing the XML document and all attachments in one zip file.
  - Although Zipped was the official standard, unzipped XML documents were eventually accepted by the Parse-and-Load software because most submissions were coming in unzipped (and because it took very little effort to support both methods).
  - If it's determined in the final implementation of OWWQX to enforce the restriction of Zip files only, we probably need to request that CDX enforce this rule for OWWQX (so that the submitter is immediately notified).
- CDX Interface
  - Interface between Parse-and-Load and CDX is one way only (calls are made to web services at CDX for everything). This reduced the need for custom programming at CDX because we relied on web services that already existed. It may have also reduced security concerns because no outside system needed access to the Parse-and-Load software.
  - One limitation is with notification of new documents at CDX. For the Pilot, the Parse-and-Load software would regularly query CDX to see if there were any new documents.
    - For the final implementation, consideration should be given to implementing a web service that CDX could call to notify the Parse-and-Load software of new documents.
- Insert, Update & Delete
  - The Parse-and-Load Software determines automatically whether a particular element is to be inserted or updated in the database. It bases this decision on whether the element's unique identifier could be found in the database. If it can be found, an update of the data is performed. Otherwise the new data is inserted.
    - An alternative would be to include an attribute on specific elements in the schema to indicate if a particular element was to be Inserted or Updated.
  - Because IDs are used to identify whether a particular element exists in the database, the granularity of updates is limited to the level in the schema that defines unique IDs. More specifically, this means that Projects, Monitoring Locations, and Activities can be updated, but Results cannot because Results do not have unique IDs in the schema. Therefore, if a Result changes in a local system, the entire Activity (and related Results) needs to be resubmitted to OWWQX to correctly keep the two systems in sync.
  - Some specific rules were implemented for updating the following elements (which do not have unique IDs): AttachedBinaryObject, OrganizationAddress, ElectronicAddress, and Telephonic
    - These elements cannot be updated individually. It is required that the complete set (relating to a specific parent element) be provided whenever updating one or more of them.
      - For example: If an Organization which currently has two addresses, needs to have one of them updated, the update submission must include both addresses. Providing only one will result in only one Address remaining for this Organization (once the submission is processed).

- In cases where none of the elements in the set have changed, none need to be provided.
  - This is done to avoid the need to resubmit this data when none of them (for a particular parent element) has changed.
  - For example: If an existing Project (which has several AttachedBinaryObjects already in the database) is included in a submission file and none of its AttachedBinaryObjects need to be updated (because they have not changed) then that Project should not include any AttachedBinaryObjects in the XML submission.
- One nuance that should be understood is that providing empty tags (i.e. tags with no data) will effectively remove the entire set from the database.
  - For example: including "<Telephonic></Telephonic>" or "<Telephonic/>" will remove all telephone numbers from the database (for the parent Organization).

Pilot Project Challenges

ESAR Data Standard

The standard was in the development mode when the pilot was initiated. This required that the WQX team freeze the standard and use what was available despite the knowledge that the standard would change.

Designing the Flow

In designing the flow, the nature of the Water Quality monitoring business process was critical. Data synchronization is accomplished in STORET by complete drop and replacement of data at the data source level. Due to increasing data volumes and cost reductions, the participants knew early on that there was a need to identify, capture and deliver changed data to the target data warehouse. It was incumbent on the data provider to determine what data had been changed.

Schematron Rules and Implementation

There was some moderate difficulty in ramping up with schematron implementation. Although the EPA/CDX team was very helpful in supporting the implementation of schematron in the CDX environment, the process would have been quicker if written documentation existed stating the requirements for the following:

- Schematron file creation
- Database-lookup protocol
- Schematron integration instructions

It is highly recommend that the EPA/CDX team develop a schematron implementation package for schematron developers. This documentation would include the CDX integration steps/protocol and potentially example code.

In addition, implementation of schematron rules that depended on database lookups used licensed 3rd party software that was not easily distributed. If CDX's schematron service

continues to be viewed as a long-term enterprise solution that supports various EPA data flows, it would be better if a schematron solution for database-dependant rules either did not depend on licensed 3rd party software, or CDX could distribute licenses without license concerns.

Data Mapping

Participants had to overcome major challenges to successfully map their data to the OWWQX schema. They needed to map their data to the OWWQX data types and effectively convert their data in a standards-based, cost-effective manner. In some instances, the semantics of data could not be inferred from syntactical clues in their representations and values, such as schema data element names, types, structures, constraints and value patterns. The participants source data were stored using different tools and methods such as the following:

- Data models and representation.
- Structural conventions.
- Naming conventions.
    - Same name used for elements and different meanings.
    - Different names for elements sharing same meaning.

Even data elements that referred to the same concept were nonequivalent, this was due to differences in unit, resolution, precision, aggregation, or measurement protocols; often hard to interpret, difficult to acquire, undocumented or unknown. The OWWQX pilot business rules were enforced by the schema.

In order to come into compliance with the EPA SoR, the EPA registry name was used when possible. However, not all water quality monitoring values needed for the pilot have been established in the SoR. Therefore, it was necessary to create and maintain a hybrid list of domain values. Research and analysis of the necessary data needed for water quality monitoring is still required to alleviate the need to maintain this hybrid list.

Even though participants overcame the data mapping challenges during the pilot phase, data partners may encounter the same challenge during implementation. It's important to note that the ESAR standard was still a working draft, work in progress. The ESAR data standard was finalized in January 2006.

OWWQX Issue Log

The OWWQX Issues Log was created to capture bugs, issues, comments and concerns encountered during the pilot lifecycle. The log was revisited often and updated as required.

Id	Title	Description	Comments
ISS8	.NET Framework at RTP	The Microsoft .NET framework shared environment is currently not supported at RTP. Should the parse-and-load routine be rewritten in JAVA or some other language?	CSC originally said a .net environment could be supported, but there was no shared environment, and no timeline for creating one, so it required separate servers. RTP is all on shared servers, no .NET environment available, hopefully summer 2006. A fallback plan is needed in case NCC/RTP is unable to support .net in a shared environment. JAVA is supported under an already existing environment that OWWQX could have access to sooner. However replacing the existing .NET code with JAVA would take considerable effort to rewrite. Java may not necessarily be the same as the pilot since it is written in a different language. .NET has been thoroughly tested through the pilot and successful, changing to JAVA at this point means it wouldn't be tested. **DECISION**: stick with .net, even if a new server is needed. Dave suggested buying a new server anyway b/c in a shared environment; users are at the mercy of other programs.
ISS9	Data Flow	For the pilot, States push data to EPA. In production, should both a pull and a push be supported?	Texas particularly wants this as an option. Push is sufficient for reporting to EPA, it is a safer option and the states are more comfortable with this option. Pull is a good idea because it is faster, computer to computer, less manual data entry/human error, but creates security risks, and will only help the powerful states since their technology will support a "Pull" option. Smaller states don't have the technology to support it, so both options (push and pull) would have to be implemented. Must keep in mind the immediate functionality, as well as long-term benefits.
ISS10	STORET Characteristics	SRS does not include all of the current STORET characteristics. Work on this is continuing to occur.	This is a very big vulnerability issue because SRS is not under OW control. Peter wants a meeting with OEI to layout what needs to be accomplished in order to support CDX. A list of priority items (characteristics) needs to be created, so it is known what must be in the database that is critical to OW work.
ISS11	Unique Activity IDs	Activity IDs must be unique for a given Organization. Will some organizations have difficulty maintaining these IDs?	This solves the problem of submitting duplicate data. Pilot participants had no problem with this, but other users will experience some difficulty; will be an adjustment, but a positive one. This issue will be worked with the user groups to get their input.
ISS12	Activity Depth	Do the current Activity Depth data elements and business rules cover user's business requirements?	This is currently required for any activity that includes samples. STORET has 20 depth fields; it will be hard to answer to why certain things aren't included in the schema and some users won't be comfortable with a reduced data set. Peter's suggestion is to ONLY include what is critical to the warehouse. Must be high-level decision on what's included or not.

Id	Title	Description	Comments
ISS13	Probability Surveys	Does the current schema support probability surveys?	The schema must include probability surveys.
ISS14	NWIS Remark Codes	ResultDetectionConditionText and/or Result Lab Comment domain values may need to be expanded to include NWIS Remark Codes	Issue tabled for later discussion.
ISS15	NWIS data	Additional MonitoringLocationType codes may need to be added to support NWIS data	Issue tabled for later discussion.
ISS16	Monitoring Location	Electronic Address information may be added to the Monitoring Location portion of the XML schema to accommodate NWIS station contact information	Issue tabled for later discussion.
ISS17	HUC	HUC should be added to the schema as an optional data element.	The states can keep using the local system, information can still be derived if needed. It will be included in the schema as optional data element. This will be an issue when discussing with the states.
ISS18	Parameter Code	USGS requested that parameter code and parameter code source be added to the XML schema	Issue tabled for later discussion.
ISS19	Required Sample Description Fields	Are there too many required Sample Description Fields for cases when Activity Type = Sample (i.e., SampleHoldingContainerMaterialCode)?	Need to scale back sample description fields, there are too many and some don't make sense for certain activity types. It gives room for people to create incorrect data that gets successfully submitted.
ISS20	Query Backend Database	Provide the ability for submitters to query the backend database to assist in deciphering the insert, update, and delete error and warning messages encountered during submissions.	This allows for status checks at any point during the process. It should be an easy addition, and will be very critical for the web tool that will be developed. Dwane wants to know how this is possible through firewalls, issues at RTP, web services, etc.?
ISS21	CDX Processing Report	Provide one processing report from CDX, i.e. provide one processing report with contains the download status and validation results with reference to the original XML submission file.	Need to discuss with CSC/CDX. EPA contact Mike Hart.
ISS22	Notification Message for Data Submitter	Provide an automated notification message to the submitter when processing is complete.	Need to discuss with CSC/CDX. EPA contact Mike Hart.
ISS23	Download Individual Documents in CDX	Provide the ability for users to download the processing report or other individual documents without having to pull all of the documents from the original submission.	Need to discuss with CSC/CDX. EPA contact Mike Hart.
ISS24	Unit of Measure Code	The business rule states that ResultValueMeasureUnitCode is "Conditional: Required if a non text result is reported." However, this rule is not enforced by the schema or schematron.	Need to correct schema to enforce the business rule.

Conclusions

This document depicted the overall technical activities during the OWWQX pilot phase and highlighted some of the practical and technical issues encountered. As OW moves into the production phase of the Water Quality Monitoring System, this document may serve as a technical reference and communication tool specifically in the following aspects:

- Updating XML Schema and Schematron files
- Updating Domain Validation Lists
- Updating ODS design and development
- Updating Parse and Load software
- Implementation of ODS production environment
- Integration of CDX and ODS production environments
- Enhanced coordination with OEI in keeping OWWQX domain values aligned with SoR

Communication was the key factor to the overall success of the pilot. There were technical difficulties inherent in the efforts of making contact with the right person, who had the needed expertise, and doing so in a timely fashion. This is not to say that the OWWQX did not have an accessible dedicated group of experts, but that the number of players corresponding to the number of parts of the projects tended to extend response time to challenges. Challenges appeared at numerous occasions, from the initial development of schema and valid values, to the development of the flow configuration document, to the testing of validation and ETL services. Participants were committed to the success of the pilot in addition to performing their day-to-day activities. The open communication channels between OEI and OW were extremely beneficial as The Office of Information Collection played a crucial role in the overall architecture of the pilot. Several communication channels were opened to foster communication among participants, such as OWWQXPILOT Quickplace, bi-weekly teleconferences, CDX help desk, and email. Despite the issues, challenges, concerns, and comments, the pilot was a huge success and was able to achieve it's objectives within the time allotted.

The results and key outcomes of the pilot activities were as follows.
- Oregon successfully submitted files via Exchange Network node.
    – Initial submission on 11/29/2005 (800 results).
    – Insert, Update and Delete functionality (~21,000 results inserted, ~9,500 results deleted).
- Wind River successfully submitted files manually.
    – Inserted ~32,000 results using CDX Integration Test Tool on 1/24/2006.
- MI completed data mapping to the schema, but did not submit.
- TX still in system development, and did not submit.

In February 2006, EPA established a Water Quality Exchange production system implementation schedule. The OWWQX pilot supports field observations, water chemistry and fish tissue information. Moving into production, OWWQX will contain the data from probability surveys that characterize condition of the nation's water resources. It also will contain data that supports measures of incremental progress towards restoration or protection of water body segments or watersheds. Web services and a XML generation tool will be available for use by data providers. The proposed schedule through January 2007 is presented in the following table.

STORET/WQX Project Schedule	
Approx. Date	**OWWQX Activity**
Jan. 2006	OWWQX Pilot Complete, begin evaluation
Feb. 2006	Begin Informatica Test Phase (rework method for loading data warehouse)
Mar. 2006	Complete OWWQX Evaluation
Apr. 2006	Begin OWWQX State/Partner Outreach (to last thru July 2006) related to the STORET vision
Apr. 2006	Begin Finalization of OWWQX Schema for Phys/Chem results

STORET/WQX Project Schedule	
Approx. Date	**OWWQX Activity**
June 2006	Begin Web Services Development
Sept. 2006	Finalize OWWQX Schema for Phys/Chem Results
Sept. 2006	Begin Pilot for OWWQX Schema for Bio/Habitat results
Sept. 2006	Complete Informatica Test Phase
Oct. 2006	Beta Web Services
Dec 2006	Data Warehouse ETL tool complete
Jan. 2007	OWWQX Schema for Phys/Chem results in production
Jan. 2007	Web Services Final (at a minimum services that serve the reference tables to support OWWQX)
Jan. 2007	Begin XML Generate Tool Development

APPENDIX - BI-WEEKLY PILOT MEETING MINUTES

OWWQX Pilot Biweekly Conference Call
Wed April 27 1-2 pm

Draft Meeting notes

- We would like this group to serve multiple purposes, including information dissemination and a forum for pilot participants to exchange experiences and lessons learned.
- Project Scope
    - Will build an ESAR data repository that will be a transactional database
    - Goal is to have pilot data flowing by the end of the year
    - After December, if we move into production with this system, then we will build the ETL software to transform the incoming data into the STORET Central Warehouse. This step is not part of the pilot, but is a follow-on after the pilot
    - The scope of the pilot is limited to water (chemical, physical, and fish tissue). Will this scope be expanded to sediment after the pilot? This will be determined based on the success of the pilot.
- The OWWQX draft schema is posted at exchangenetwork.net. A Flow configuration document is to follow (M. Hart will follow up)
- Expectations:
    - **Texas:** currently working on a new data management system to increase capabilities to provide data to EPA. Not be able to participate in the pilot to submit through a node in that fashion. But they will participate.
    - **Oregon:** database is production ready database. Have a node running. Will expect to deliver submissions through CDX. As soon as flow configuration document is defined, they will be able to map and get their submission ready. They are particularly interested in dealing with updates / deletes issues. They plan on including fish tissue samples. They also include all raw continuous monitoring data as part of their data. They could go either way on that data. They have 2 databases: Pacific Northwest (that includes data from other NW states), as well as their own OR database. Database goes back to 1945. A nearly complete overlap of legacy STORET and OR database.
    They expect that the schema may need to go through modifications as we go through the pilot.
    - **Wind River Indian Reservation:** In the middle of doing a data management system update. Most data is water quality data, will at least try to prepare water chemistry at the same time they prepare their air data. They are currently getting their Node up and running. They have about 10 years of water quality monitoring data. Need 6 weeks of programming time to get their node up and running.
    - **Michigan:** Michigan had to leave the call early. They will have the opportunity to discuss their expectations in more detail at the next meeting. Jason Smith will be Michigan's representative.

- Additional Questions
    - How do we avoid potential redundancy of data that is submitted? This is not a major concern for the pilot but a plan needs to be developed before implementing a production system. This issue will be discussed with Oregon in more detail.
    - What about making this data back available to states by publishing on the CDX node? During the pilot, there will be some services to retrieve back the submission you had made. Longer term is to build some data retrieval services to grab data from data warehouse.
- Action Item - An email distribution list of OWWQX stakeholders should be established. J Wilson will follow-up.
- Our next call will be on 5/11 from 1:00-2:00 EST

Oregon DEQ Pacific Northwest Data Exchange Lessons Learned
--Focusing on Create, Replace, Update and Delete (CRUD) Issues
Call Date: May 2, 2005
Time: 11:00 a.m. to 12:00 pm.
Location: Teleconference hosted by Lockheed Martin

_____

Mitch West (OR), Glen Carr (OR), Curtis Cude (OR), Cary McElhinney (EPA), Joe Wilson (EPA), Marty McComb (EPA), Lee Manning (EPA), Lynn Singleton (LMIT), Marybeth Puckace (LMIT), Douglas Timms (EnfoTech), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems)

_____

Pacific NorthWest Data Exchange Network "CRUD Lessons Learned

Background: The Pacific Northwest Data Exchange Project has an established Exchange Network Node and has been managing/exchanging environmental information with the others in the PNW for some time. The PNW Data Exchange Network offers a service to those who want to post their data and do not have a node.

ODEQ agreed to share some of their Lessons Learned with the ESAR Pilot Team.

M. West reported some of their observations, concerns and issues which might be considered for the upcoming ESAR pilot program.

Transaction based system relies on the data provider to determine what has been changed. Replaces data in chunks versus record by record. The later requires more work, more complex.

The data sets are highly portable and they can be large files; however they compress well into ZIP files.

- Environmental data sets are relatively static
- They must rely on the data stewards to manage the data quality and be responsible for knowing what they have submitted
- Duplications are the data stewards responsibility to prevent
- Create, Update, and Delete functions are at the heart of their success and several operational and governance considerations were needed.
- Creates are handled in two areas—Project and Station and Results
- Wholesale delete and replacement with each submittal is not practical due to file sizes
- Similarly, the overhead associated with record by record update is not practical

As such, they have segregated their CRUD updates into 5 types:

1) Project and Station Only—used to establish—a unique ID is required to track results
2) Results only—posted to an established Project and Station—added information.
3) Results only delete and replace
4) Wholesale replace for Project, Station and Results
5) Record by record replace

Items 1-4 are handled automatically and instructions are noted in the headers of the submittal e.g., Delete cascade, Delete Transaction, Change Transaction etc. .

Item 5 is rarely done and handled on the phone.

Governance—Trading Partner Agreements are needed to specify any governance issues and the details of the exchanges. These also address the submittal of duplicate data or other operational norms.

PNW Data Exchange allows for Project and Stations to be grouped and identified with an unique identifier. The current Pilot schema does not have this option present. The ESAR standard does allow for these identifiers.

Need to further refine the replace transaction and stated in the past that the data provider would make a phone call in order to complete a change/deleted transaction; however this could be a problem once this system is implemented on a National level.

PILOT SCOPE

The proposed pilot was also discussed during the call. The following topics should be considered:

The pilot will test functionality

Need to define data file sizes

Data sets at the project level may be a good place to start

Need to submit new data from multiple projects

Refine the governance issues—the data provider will be responsible to ensure that duplication does not occur.

The Transaction data sets need to be defined in the flow configuration document

Pilot needs to have some way for the data steward to determine/evaluate what data has entered the warehouse

Update categories need to be defined—as noted above or other options

Update frequency

Granularity of the data sets to be used in the pilots needs to be determined

QA/QC issues need to be defined

Feedback forms following submittal should be considered

Web Sim feeds require manual entry of Project and Station—this could be a factor for a full scale implementation

OWWQX Logical Data Model Discussion

Call Date: May 23, 2005
Time: 3:00 pm. to 3:45 pm.
Teleconference hosted by Lockheed Martin

Lee Manning (EPA), Marybeth Puckace (LMIT), Douglas Timms (EnfoTech), Dave Wilcox (Gold Systems), Ryan Jorsengen (Gold Systems)

Background: The OWWQX Logical Data Model was presented at the May 13, 2005 meeting. This was a first attempt to develop the data model based off the XML Schema. Review of the Data Model by EPA resulted in several comments. This telecom was scheduled to further investigate those comments.

Discussion:

1. The relationship between subsample/sample, result and activity needs to be better defined; the data model correctly reflects the schema. Need to resolve scenario- A sample can have many subsamples with many results for different laboratories, this is currently not reflected in either the data model or the schema.

2. State Table, a little confusion between postal code one meaning state ID(PA) vs Zipcode. EPA states that there is a state code table available on its website. Need to model accordingly and align with data registries.

3. Organization and Co-Operating Organizations how to best represent this relationship, it's ownership and Organizations/Co-Operating Organization ID's. Suggestions include: a free form text in the Organization box.

Action Items:

Gold Systems to revise data model to best represent item #1and#2. Keeping in mind the following:

Database Design and least impact/minimal change to the XML schema. Gold to communicate with enfoTech and how the new Data Model could impact the Schema.

Check out the EPA Website –State Table

Try to have next data model draft to Lee Manning (EPA) by June 1st.

OWWX Pilot Biweekly Conference Call
Call Date: June 1, 2005
Time: 1:00 a.m. to 2:00 pm.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Lee Manning (EPA), Cary McElhinney (EPA), Mike Hart (EPA), J.J. Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems)

The purpose of today's meeting was to meet with the OWWQX Pilot Team and review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- This is the last meeting that L. Manning will be attending due to his retirement.

- K. Gunthardt is a new member of the monitoring branch.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

- The data model is being developed in Oracle ERG based on the XML schema. The team should have a work product to share in a couple of weeks.

- The Flow Configuration Document (FCD) was added to the task. No schedule has been developed for the project yet. M. Hart indicated that by the end of next week a working template should be developed. He added that due to unforeseen project constraints the document could change slightly over time.

- A sample data submission document which includes live data value examples was developed in XML to complement the data dictionary. This document will be tested and then sent out for review before the next meeting. Participants can review and comment on the draft schema located at exchangenet.net.

DATA TRANSACTION RULES AND LESSONS LEARNED FROM OREGON

- Minutes were distributed to Pilot Participants on Lessons Learned from the PNW Data Exchange Network.

- The PNW Data Exchange Network team stated, future data transactions, will be difficult to determine what submissions are modifications of old data versus entirely new data submissions.

- J. Wilson suggested that data such as organization, project, monitoring location, and activity information should be maintained by the system. He added that the system should have the appropriate level of granularity so that each submission does not have to be a complete drop and replace. He proposed that the system use the Activity as the smallest level. An activity ID would exist for each location.

- The participants agreed that consistent activity IDs should be maintained from one submission to the next.

- One participant suggested that the data submitters could use the header scheme of the XML document to specify whether a submission is new data or an update of existing data.

- In addition, the Activity ID could be checked against the existing objects to see if the object already exists.

- One concern was how to indicate when a record should be removed. In the future the system will have to be able to support larger delete IDs.

Action Item: A standard header scheme from FRS will be distributed before the next meeting along with a list of pros and cons of usage of the header schema approach.

Central Data Exchange Water Quality Monitoring Flow

System Requirements Specification (SRS) Document

Action Item: Define the requirements for the CDX portion of the project and submit the final draft.

- To avoid confusion OWWQX will be used except when referring to the pre-existing ESAR Data Standard.

- For the purpose of the pilot, the draft requirements document specifies a limit on the file size of submissions. If necessary, that file size limit can be modified for the final production version of the system.

- C. Cude noted that section *2.14 Miscellaneous* part *2.14.2.1* indicates that the system will be available during normal working hours (8 a.m. to 8 p.m. [TBR] Eastern Time). He suggested that the system availability should be extended.

- Section *3.6 Use Case 5: File Verification using the QA Server* states that the submission of the file and all verification occurs first. He indicated that a precondition for Use Cases 1 and 4 is that verification and schematron check occur first. These will refer to Use Case 5. Updates will be made to the version of the SRS submitted by M. Hart.

NEXT STEPS

- The draft data model will be distributed for review.

- The timeline will be added to the Flow Configuration Document (FCD).

The next meeting will be held on June 15, 2005.

OWWX Pilot Biweekly Conference Call
Call Date: June 15, 2005
Time: 1:00 p.m. to 1:30 pm.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Lee Manning (EPA), Cary McElhinney (EPA), Mike Hart (EPA), J.J. Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems)

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

OWQQX Logical Data Model/Flow Configuration Document

- The Logical Data Model was distributed to Pilot Participants prior to the teleconference for their review. Dave Wilcox (Gold Systems) gave a brief overview of the ERD and spoke about some of the key relationships.

- A coordination meeting between CDX and LM will be held to reach an agreement on how the data handshakes are going to occur. RTP will define the hosting requirements.

- M.Hart stated the Flow Configuration Document would be available Friday, June 24, 2005.

Exchange Network Document Header Specification

The Header Template will have an operations tab to enable specific key usage for data access. The question was raised, how will it work and will its schema be compliant to the standard. A discussion occurred on occasional variances that might occur. This was mostly tabled since the discussion of standards should be in a different venue.

Central Data Exchange Water Quality Monitoring Flow

System Requirements Specification (SRS) Document

- A revised SRS document was distributed to pilot participants prior to the teleconference. The revised SRS document corrected items, which were noted in the June 1, 2005 OWQQX bi-weekly teleconference.

    OWWQX Pilot replaced references to ESAR Pilot.

    Sections 3.6 Use Case 1 and 4 was changed to reflect preconditions found in Use Case 5 that all file submissions are check via the QA server for Schema and Schematron violations prior to submission to the OWWQX Pilot System.

Next Steps

Review Flow Configuration Document

Additional agenda items please forward them to Joe before the next meeting.

The next meeting will be held on June 29, 2005.

OWWX Pilot Biweekly Conference Call
Call Date: June 29, 2005
Time: 1:00 p.m. to 1:45 pm.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Lee Manning (EPA), Cary McElhinney (EPA), Mike Hart (EPA), J.J. Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River) Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems)

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

OWQQX Flow Configuration Document

- The purpose of the Flow Configuration Document (FCD) was described by Mich West (OR). The FCD will outline the types of data that can be submitted, how this data should be formatted, and how the users will interface with CDX to submit their data. The FCD will not include any backend communication or processing information.

- The Draft FCD was distributed to Pilot Participants prior to the teleconference for their review.

- The Draft FCD still needs to be developed to address submission types- Insert/Update, Delete, and Submittal Status Requests. A sub-group was formed to discuss the FDC in detail, members; Joe Wilson(EPA), Mike Hart(EPA), Mitch West(OR),Curtis Cude(OR), Dave Wilcox (Gold), Ryan Jorgensen (Gold), Natalia Vainshein(CSC/CDX)

Other Notes

Curtis Cude (OR) addressed the issue of how binary objects should be delivered to the CDX. The current plan is to have the binary objects delivered external to the XML and reference from within. It is believed that method this should comply with all current standards.

Mike Hart (EPA) stated and pilot members agreed that there is little value in developing data retrieval capabilities against the OWWQX database. Once the data has been loaded from OWWQX into the central STORET warehouse, participants will be able to pull their data from this location.

For the purpose of the pilot, it was agreed that the pilot participants will submit or 'Push' their data to CDX. It was discussed, however, that the final implementation may support an option where CDX may solicit, or' Pull' data from the participant nodes. The FCD will initially only outline the Push procedures available in the pilot.

Next Steps

Curtis Cude (OR) will distribute the Pacific NW WQ Data Exchange FCD to Pilot Participants.

Sub-group teleconference call to discuss FCD is scheduled for Thursday 30 from 12-2pm, CSC will send out dial-in information.

Discuss recent changes made to the Schema.

The next meeting will be held on July 13, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: July 13, 2005
Time: 1:00 p.m. to 1:30 pm.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

OWWQX Flow Configuration Document

- A revised draft Flow Configuration Document (FCD) has been distributed to the FCD Sub group for review. CSC and the FCD sub-group have made significant process and a new draft FCD will be ready and distributed to pilot participant prior to the next OWWQX Pilot teleconference. .

- Off-line discussion will occur on the timing of data submissions.

OWWQX Schema/ Schematron

Doug Timms (Enfotech) gave a brief overview on the minor changes made to the Schema**.**

- Delete Schema needs to be developed

- Activity Organization field is now a free-text field instead of an ID field.

- Changes to the way in which subsamples are handled in the schema

- Slight reorganization to the files

- More changes may occur based on how the final OWWQX FCD document will look, this will include the development of a Delete schema.

Other Notes

It was stated that both the FCD and Schema are evolving documents and will likely change as we work through the pilot. It was also noted the need for good version control on documentation to minimize duplication of efforts.

Next Steps/Discussion

Deployment of QA services; discuss how data submission can be validated against the Schematron.

Joe Wilson will provide participants a package to include: Domain Spreadsheet, FCD, Schema/Schematron

The next teleconference will be held on July 27, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: July 27, 2005
Time: 1:00 p.m. to 1:40pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

General Comments

- Joe stated steps have been taken internally to broadcast the OWWQX Pilot. Also, the pilot is receiving a lot attention from HQ.
- Initial steps in identifying tools (Informatica) for mapping the OWWQX data into the STORET Central DWH.

OWWQX FCD

- The latest distributed version of the FCD included input from both Oregon and Gold Systems. For the purpose of the pilot it was decided not to "Append" results, instead it would be drop and replace. Also, if an activity relates to multiple projects and one of the projects are being deleted the Activity would not deleted.
- CSC stated Windsor Solutions will perform a 3rd party review on the FCD prior to distributing.
- Delete Schema has been created.

QA Services

- A QA test file will be developed by Mitch West.
- Schematron to be updated based OWWQX ODS.
- For the pilot, OWWQX lookup tables will exist in the CDX environment as well as the Central Environment. Methods for keeping these tables synchronized will be address for production.

Next Steps/Discussion

- QA Services-Testing Plan

- OWWQX FCD

- Distribute to pilot participants an Access DB Version of the OWWQX database/reference table values.

The next teleconference will be held on August 10, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: August 10, 2005
Time: 1:00 p.m. to 1:45pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products.

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, made announcements, and reviewed the topics on today's agenda.

- Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

Updated Work Projects

- Delete Schema. After the June 24th distribution, no functional changes have been made to the schema. The delete schema was incorporated into the overall schema.

- Sample XML files were distributed for testing purposes.

- The Schematron files were updated for the QA Services business logic on the CDX end. The files now point to the new look-up table structure.

- The Excel spreadsheet that documents the different elements, their definitions, and whether they are required, was updated to include the rule ID for each element.

  o The root element of the delete schema and the root element of the update schema are different files that share the same namespace in order to accommodate the URL database. Doug agreed to change the root file to point to the other root elements.
  o In addition Doug will create an instance document for the Schematron for all of the rules for testing.
  o In response to a question Doug indicated that it is necessary that some Schematron rules to use the database lookup and some do not in order to accommodate external users.
  o Dan of Wind River noted that one issue with the Schematron is that there is only room for one Tribal Code. This is a problem in the instance when a program is a joint effort between 2 Tribes.
  o Contact Mike Hart at hart.michael@epa.gov or 202-566-1696 in reference to any issues with the Schematron.

Reference Table Values

- The original reference table spreadsheet was converted into an Access database for ease of use.

  o A participant from CDX commented that the field name in the Access table (e.g., county code) is different from that in the Oracle table. If it can be made consistent, CDX will use the Access rather than the Oracle version.

- Joe indicated that he will re-visit the issue of reference table values when the regions start mapping files to ensure that the necessary reference values are available to accurately represent the data. Where possible the reference table values are based on the EPA Registries; those values will be updated periodically.

QA Service Deployment

- Joe indicated that the QA service provides a web service where users can validate files prior to submission to CDX. Mike added that even after the files are validated through the Schematron, they will be re-validated at CDX.

- Natalia indicated that the design documentation will be ready for formal review by Friday.

Outstanding Schema Issues

Joe described the 2 outstanding schema issues as follows:

1. Activity Identifier is defined in the data dictionary as *unique for a location on a given date*. However, it can also be defined as *unique across an organization at a given location*. If the Activity Identifier is not unique for an organization, the system may not recognize entries as replacements. In this case, the user would have to submit a delete record followed by a replacement record. Alternatively, if the Activity Identifier was made to be universal, the software could distinguish between new and replacement entries; however, this requires that the Regions supply unique Activity Identifiers.

   o Marty of Region 8 indicated that the universal identifier is preferable but that the Agency does not have the resources to provide them.
   o Wind River indicated that they would be able to supply the universal identifier, but that if that solution is not used, the process should be updated to be a delete and replace each time.
   o Texas indicated that they would like to discuss the issue offline.

   Action: Joe took the action to arrange a meeting with Julie (Texas) to discuss the issue next week.

   Next Steps: Tim indicated that they will keep the current data dictionary definition of Activity Identifier and continue to investigate the feasibility of a universal identifier.

2. Binary Objects. Attached files were originally embedded in the XML but are now referenced with pointers. The issue is whether the Binary Object tag should remain in the XML. Doug recommended that it be left in the XML to allow for future use. In addition, this will keep it in line with the Shared Schema Components (SSC). He suggested that this issue be recorded in the Flow Configuration documentation.

The next teleconference will be held on August 24, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: August 24, 2005
Time: 1:00 p.m. to 1:45pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

WELCOME AND AGENDA REVIEW

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products. J. Wilson welcomed the participants, and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

Full Configuration Document (FCD) Update

- Windsor Solutions(Bill) discussed the results of his review of the FCD. In addition to some comments made in regard to readability, he had the following questions/comments:

  1. Is the package for the web-based submissions and Exchange Network submissions the same?
  2. Implementation of the header: Name value pairs will not be used for this exchange.
  3. When submissions are made, the transaction ID method 'submit' is used. An array of documents can be submitted as well.
  4. The Exchange Network Header should contain only one node.

  5. The differences between Update/Insert submissions and Delete submissions need to be clarified.
  6. Submission processing and feedback: The document should include an explanation of how feedback is returned to the submitter. In addition it should elaborate on the summary/error report that is retrievable by the submitter; schema should be developed to define the look of the report.

Activity ID

- Activity Identifier is defined in the data dictionary as unique within an organization for a location on a given date. However, in implementation it has been found that data maintenance is easier if the Activity ID is simply unique across an organization.

- Gold Systems (D. Wilcox) explained that to correctly modify a date or location field a user must submit separate delete and insert activities. Otherwise, a modification will result in the insertion of a new activity. To remedy this, it was proposed that a Universal ID be created.

- J. Wilson indicated that this solution will be implemented unless it will exclude potential users or create other difficulties. Participants from Michigan, Wind River, Texas, and Region 8 agreed that for the purpose of the pilot they have no concerns with the use of a Universal Activity ID. J. Wilson indicated that they will proceed with the Universal ID but added that he should be notified should any objections arise.

Reference Table Values

Pilot participants were asked if they noticed any gaps in the Access database of reference table values. There were no comments on the database. J. Wilson indicated that he will continue to re-visit the issue to ensure that the necessary reference values are available to accurately represent the data.

ESAR Data Standard

C. Cude indicated that the ESAR Data Standard was posted in the Federal Registry for public review and comment at the beginning of August. The ESAR Data Standard can be access through the EDSC website. Comments are due by mid November. The Data Standards team will then reconvene to make the necessary changes to the standard by the end of 2005.

QA Service Deployment

XML document validation service for the exchange network has been deployed. Those users who already have an account can access the service at http://tools.epacdxnode.net

To ask for an account on NAAS, users should send a request to nodehelpdesk@csc.com.The validations are automatically generated if the document is less that 100 MB; if the document is more that 100 MB it will be validated offline and the results will be emailed to the submitter.

Version Control

M.Hart is investigating a potential place to post the latest versions of the schema and the FCD. C. Clark will look into a central location on the epa.gov site for posting of schemas. It was noted, once changes have occurred and posted, participants will need to be notified.

State Updates

- Alaska indicated they have install SIM Web services and are interested in reviewing documentation on OWWQX. J. Wilson suggested the documentation be provided to them with the caveat that this is a pilot and therefore the schema may change.

- California is trying to coordinate the various State agencies but is interested in OWWQX as well.

Action Items

- J. Wilson will distribute a clean copy of the FCD to the participants before the end of the week.

The next teleconference will be held on September 7, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: September 7, 2005
Time: 1:00 p.m. to 1:30pm
Location: Teleconference hosted by Mike Hart (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

WELCOME AND AGENDA REVIEW

The purpose of today's meeting was to meet with the OWWQX Pilot Team to review their progress and products. J. Wilson welcomed the participants, and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

Full Configuration Document (FCD) Update

Mike will distribute the current version of the FCD for review, at the end of this conference call.

Sample/Sub-Sample Schema

Gold Systems gave an overview on the sample and subsample relationship issue, basically duplicated Sampling Method, Transport, and Storage fields should be completed for either a sample or the related sub-samples but not both. Gold distribute two alternate views for handling this issue and ask participants to review and provide comments.

Q&A services

No comments

Reference Table Values

Pilot participants were asked if they noticed any gaps in the Access database/reference table values. Julie from Texas is waiting to confirm values, intends to have the review completed with comments by the next meeting call. There were no other comments.

CDX-RTP Communication

A teleconference between Gold Systems/LM and CSC was scheduled for September 9, 2005 to smooth out any web services details and to discuss the Interface Control document(ICD), which Gold Systems will need from CSC to complete the Parse and Load Module.

Version Control

M.Hart is investigating a potential place to post the latest versions of the schema and the FCD. One idea is to post on the Exchange network.

Conclusion

Call ended with a request for any agenda items from anyone to please get them to Joe before the next meeting.

The next teleconference will be held on September 21, 2005

OWWQX Pilot Biweekly Conference Call
Call Date: September 21, 2005
Time: 1:00 p.m. to 2:00pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES

Review of Proposed Schema Changes

▪ The parse and load software should be finished by the end of September.

▪ One main issue with the schema structure is that information for submissions originally had to be stored with both the sample and sub-sample. The best option for changing the schema is to take the current elements in both places and group them with the central sample preparation and collection procedure table. Users will then have the option to fill in the sample preparation area without filling in a sub-sample. An email describing the solution was distributed.

▪ Joe indicated that an updated version of the schema will be distributed within the next couple of weeks.

▪ Multiple organizations per submission. After some discussion it was decided that for the purpose of the pilot, users will be held to one submission per organization. An element that will allow multiple organizations per submission will be added during implementation.

▪ Conducting organizations. It was decided that the one-to-many relationship for input of conducting organizations is cleaner for retrieval purposes. Multiple conducting organizations per activity will be allowed.

Handling of File Attachments in Data Submissions

▪ Attached binary objects. It is undesirable to have to include elements that do not change often (project and monitoring location) in each submission. The proposed solution was that users should provide/exclude the following data for each submission depending on the updates they would like:

a. Submit no tags and no data = No update

b. Submit empty tags = Delete corresponding attachment(s)

c. Submit tags and data = Delete the existing data and replace it with the corresponding attachments

▪ Doug Timms took the action to update the valid value list of data elements.

▪ Participants indicated that they can develop and share sample XML schemas to use to test the software by mid-October. On Joe's suggestion, Michigan indicated that they can submit blobs as a part of their pilot submission.

Reference Table Values

▪ Due to time constraints, Julie from Texas agreed to address her proposed changes to the Reference Table Values at the next meeting.

Version Control

Mike Hart found a central location on the epa.gov site where the latest versions of work products can be posted. The site will be shared with ECOS, and contractors can post to the site through email requests.

Next Steps

- Due to time constraints, Curtis agreed to discuss some changes to the monitoring geospatial location at the next meeting.

- The security plan has been developed for the application at RTP to establish a presence on the EPA servers.

- Gold and Lockheed Martin will be on the ADC call next Tuesday.

- CSC and Gold are scheduled to meet to discuss the back end system.

- Participants should send suggested agenda items to Joe before the next meeting.

The next teleconference will be held on October 5, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: October 5, 2005
Time: 1:00 p.m. to 2:00pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES
Review of Proposed Schema Changes

- Beta of the parse and load software is ready for testing. In addition, the major components of the data flow are established. The next step is to test the software with the XML data and integrate all components of the data flow.

- Doug Timms reviewed the changes made to fine tune the schema. Some of the required tags are now optional. For example, sample preparation and lat/long horizontal reference code are no longer required.

- The Access database captures what is in the Oracle database. The Excel spreadsheet should be used as a data dictionary.

Trailing Blanks and Periods in Reference Tables

The original lookup table values were taken spreadsheet format which resulted in some trailing blanks and periods in the reference tables. All participants agreed that these should be removed for the sake of consistency.

Reference Table Values

Characteristic Table Values and SRS. Joe indicated that the goal of STORET and the pilot is to conform to SRS. One participant suggested that there should be a process for keeping reference table values in sync with those in SRS. In addition, he said that there should be a way to keep track of new chemical names that are not yet EPA approved. Joe indicated that when terms do not exist in the name list, the hierarchy of characteristic names is as follows:

- EPA approved registry names
- CAS names (names that are not present in the registry)
- STORET names that are important for monitoring, but are not in SRS

Action: Joe took the action to find and distribute a list of STORET names that are not in SRS. Comments on the FCD

Action: Mike Hart took the action to find out if the FCD has been updated to reflect the Global Identifier changes.

Comments from Texas on Schema and Reference Tables
The OWWQX_TRANSACTION_STATUS Reference Table status codes

Some messages in the FCD are not in the transaction status table. Joe said that the table is internal to the data model.

OWWQX_RESULT_LAB_COMMENT Reference Table values
Joe Wilson said that extra values can be added as long as they don't conflict with existing values. Joe will distribute definitions for the existing values and then review the suggested additions.

Add surface water to the Activity Media Subdivision Table
Participants requested the addition of surface water to differentiate it from other kinds of water. The issue will be re-visited at a later date.

Split Soil/Sediment
It was agreed that soil/sediment will be split into separate column values in the Activity Media table.

Comments from Oregon on schema and reference tables
Pick lists

The pick lists in the workbook are out of sync with the Access Database. Only the data dictionary worksheet in the workbook should be used. The updated Access database will be redistributed after the meeting.

Project Identifier Field

Participants noted that the Project Identifier field is only 8 characters long. It will be lengthened to 35 characters.

Monitoring Location Type Name

A participant asked if a subset of land could be added. Other Surface Water and Other Ground Water will be added.

Source Map Scale Number

Source Map Scale Number is not a required data element in the Pacific Northwest, but it is required in OWWQX if HorizontalCollectionMethodCode. Oregon agreed to map to the worst case (1:100,000) when the value is unknown. This will be captured in the FCD.

Monitoring location geospatial elements

A participant asked if monitoring location geospatial elements that can be derived from lat/long should be required or if it can be left blank. State, County, and Country Code will be made optional for the pilot.

Next Steps

- Wind River's comments on the schema and reference tables will be addressed on the next call.

- Joe asked that all participants register for the Web page for the pilot.

The next teleconference will be held on October 19, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: October 19, 2005
Time: 1:00 p.m. to 2:00pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Natalia Vainshein(CSC), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES
Recent Updates

- CSC will provide a place holder in the Flow Configuration Document (FCD) for someone to provide text on the insert and delete operations.

- An updated copy of the Interface Control Document (ICD) will be made available to the team as well as posted on the WQX quickplace website. Some of the team members expressed difficulty in accessing the quickplace website.

- Joe stated that the updated copy of the schema and schematron with associated documentation describing the changes was distributed to the team.

- Joe will look into the suggested change in spelling for Detection Quantitation Level

- In response to Oregon's comment about Activity Type Codes, composite sample with parents and sample depletion replicate will be added. In addition, sample will be broken down into sample routine and sample other.

Reference Table Values

- Characteristic Table Values and SRS. Joe indicated that in order to come into compliance with the EPA SoR, the EPA registry name should be used when possible., Not all water quality monitoring values, however, have been established in the SoR. In the case that there is no EPA approved chemical name available, the SRS substance name should be used. In order to internally track what value is used, a CHR_SRS_ID column was added to the access database that was distributed yesterday.

Test Environment Deployment
The team is currently testing the interface between the CDX environment with the parse and load software. It is on schedule to be tested by the pilot participants next week.

Next Steps

- Joe took the action to look into the problems with access to the Quickplace website.

The next teleconference will be held on October 26, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: November 9, 2005
Time: 1:00 p.m. to 2:00pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants, introduced Dwayne Young (EPA), and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES
Recent Updates

- A meeting was held at the United States Geological Survey (USGS) to discuss using the WQX schema. Overall the meeting was positive with approximately six issues in addition to a request for optional data elements.

- The WQX pilot will contain a hybrid characteristic name list.

- Quickplace library section is being used for version control. It is here that interested parties should be able to find the most up-to-date product information.

- Issue with the spelling for Detection Quantitation Level will be tabled until after the completion of the pilot.

Reference Table Values
- The Reference Table values have been updated since last meeting. The changes are as follows.
  - ACTIVITY_TYPE table was updated (*Depletion Replicate* and *Composite Sample with Parents* were added; and Sample was replaced with *Sample – Routine* and *Sample – Other*).
  - Additional values were added to RESULT_COMMENT table.
  - HORIZONTAL_COLLECTION_METHOD was updated to use the shorter FRS codes.

- RESULT_LAB_COMMENT will be part of the WQX implementation phase.

Test Environment Deployment
The team is currently testing the interface between the CDX environment with the parse and load software. The issue of getting the processing report to CDX should be resolved by close of business today.

The pilot is not expected to be in compliance with the Cross-Media Electronic Reporting Rule (CROMERR).

The backend software is currently being hosted at Gold Systems. The software will be moved to the CDX environment over the next few weeks. This should be a seamless transition process.

Next Steps
Curtis will submit a few transactions to make sure data is flowing as expected. The WQX flow will remain in the CDX pre-prod environment for the duration of the pilot phase.

The next teleconference will be held on November 16, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: November 16, 2005
Time: 1:00 p.m. to 2:00pm
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Dwane Young(EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

*Due to the Thanksgiving holiday, the next meeting will be held on **Wednesday, November 30th**.

STATUS UPDATE AND OWWQX ACTIVITIES
Initial Testing of QA Services and the CDX Data Flow

The following issues were encountered:

Org IDs: organizations need to be registered before submitting data. For the purposes of the pilot, users should create a new 8 character Org ID rather than using their existing STORET ID.

Joe took the action to confirm that the Org IDs and Beaches IDs do not conflict.

Schematron Testing: an issue arose having to do with the validation of georeference items

Sample Preparation Business Rule: Sample Preparation category requires that the sample description be provided if the activity type includes the word "sample". In reality, sample description is not required if it is not a sample (e.g., field measurements observation). It was noted that some Quality Control activity types are field measurements that might be samples but do not contain the term "sample" in their name. Joe took the action to add the word sample in such instances.

Result Detection Condition Text enumerated list: the enumerated list embedded in the XML schema will not allow a null value. Option 1: change the Result Detection Condition Text enumerated list to allow a value within the quantitation limit. Option 2: drop Result Detection Condition Text down one level. Workaround: Do not use the enumerated list and handle the text as a string; this option could cause problems in implementation. Doug suggested that the null string in the enumerated list be moved to the top of the list to serve as the default.

Characteristics Names List: Joe indicated that the latest list that was distributed is missing physical characteristics in STORET. The list will be updated with the STORET characteristics names by the end of the week.

Activity Depth Altitude Measures: Joe took the action to confirm that the logic is the same as that for the Sample Preparation Business Rule.

User Support

- The CDX Helpdesk can be reached by calling 1-888-890-1995 and selecting option 2 for the node helpdesk.

Current Status of Testing

- Oregon is currently doing some initial testing. Michigan is on target for testing at the end of December. Wind River is on target for XML generation but the node may not be in place onsite by the end of December; if not, an offsite node will be used for testing.

Procedural Guidance

Joe indicated that the Flow Configuration Document (FCD), schema, Schematron, and other technical documentation with be updated post-pilot. In addition the first draft of the Exchange Network Flows Lessons Learned document will be generated next week.

The next teleconference will be held on November 30, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: November 30, 2005
Time: 1:00 p.m. to 1:30 p.m.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Dwane Young(EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME AND AGENDA REVIEW

J. Wilson welcomed the participants and reviewed the topics on today's agenda. Participants who would like to be added the distribution list for this meeting should email Joe Wilson at wilson.joe@epa.gov.

STATUS UPDATE AND OWWQX ACTIVITIES
Business Rules

Sample Preparation Business Rule: If the activity type contains the text "sample" then you must provide a discrete activity depth or a range of activity depths.

If the activity type is sample integrated vertical then you must use an activity depth with top and bottom designators.

Doug took the action to identify the non-database dependant rules within the data dictionary spreadsheet.

Data Submissions Update

- Oregon successfully completed their initial submission (800k) which took a few seconds to load.

- Curtis took the action to look into the namespace/header issues.

- Oregon will test a larger data submission later today.

- Michigan is still on track for data submission by the end of December.

User Support

The CDX Helpdesk can be reached by calling 1-888-890-1995 and selecting option 2 for the node helpdesk.

Joe added that he can be contacted in regard to user support issues as well.

Lockheed Martin is compiling a list of issues reported throughout the pilot. The draft of this list will be distributed before the next call.

Joe took the action to ensure that Doug's schema update is distributed to all of the participants.

The next teleconference will be held on December 14, 2005.

OWWQX Pilot Biweekly Conference Call
Call Date: December 14, 2005
Time: 1:00 p.m. to 1:30 p.m.
Location: Teleconference hosted by Joe Wilson (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristin Gunthardt (EPA), Dwane Young(EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME AND AGENDA REVIEW

Joe Wilson reported that he has now transferred to the EPA OW front office but said that he will still be working with the STORET team and with the Pilot. His new telephone number is 202.564.2867. Kristin Gunthardt will be taking over the OWWQX Pilot activities.

Kristin welcomed the participants and reviewed the topics on today's agenda.

STATUS UPDATE AND OWWQX ACTIVITIES

The updated Data Dictionary (with additional STORET mapping fields) is available on QuickPlace.

Data Submissions

- Curtis indicated that the Schema header issue was resolved.

- It was reported that the asynchronous response issue was due to Schematron server problems. There is still an issue with the response email.

- The delete submission went through successfully.

Issues Log

- Kristin indicated that the draft Issues log that she distributed yesterday will eventually be included in the Lessons Learned document at the conclusion of the pilot.

- Joe reviewed the status of a number of comments on the schema submitted by USGS:

  o **.net vs. java:** Currently there is no shared environment for .net and no deadline for creation of such an environment. This means that a separate server would have to be purchased.

  o **Parameter Code Field:** NWIS currently rolls a number of characteristics into a numeric code. They indicated that they are willing to map the component pieces of that code to the schema; however, they requested that an option field be created as a placeholder for that code.

  o **NWIS/USGS Interaction with SRS:** Maintenance of and mapping to SRS was a major topic of discussion.

  o **Legacy Remark Codes:** this issue is being resolved.

  o **HUC field:** may need to add to the schema as an optional field.

  o **Contact Info at the Monitoring Location Level:** proposed

- Joe provided the following highlights from the morning meeting with USGS:

  o Ken Lapierre is transferring his responsibilities as NWIS coordinator to John Scott.

---

      o   Due to their concerns about funding, USGS said that it will take them about one year to map their data to the schema.

Next Steps

- Curtis took the action to provide an update submission.

- Kristin indicated that although the goal is to conclude the pilot at the end of December, the testing phase can be extended to mid-January to allow one of the other participants to submit data. She will follow up with Michigan, Wind River, and Texas to determine whether an extension will be beneficial.

- By Monday, the issue regarding submission of zip files with multiple attachments should be resolved.

The date of the next teleconference will be announced via email.

OWWQX Pilot Biweekly Conference Call

Call Date: January 11, 2006
Time: 1:00 p.m. to 2:00 p.m.
Location: Teleconference hosted by Kristen Gunthardt (EPA)

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristen Gunthardt (EPA), Dwane Young(EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

WELCOME

Peter Grevatt, Monitoring Branch Chief thanked the states, tribes, contractors, and the EPA team members for participating in the pilot. He stated this pilot was the beginning of exploring new ways for interested parties to share water quality data with EPA . Currently interested parties need a local copy of STORET in order to exchange water quality data with the EPA. The pilot began to explore how to flow data using the EPA exchange network (CDX) and a schema based on the ESAR Standard. Due to the successful submission of data by the PNW the pilot will be brought to an official close.

In the next stage, EPA hopes to bring in a larger set of partners to explore the lessons learned from this pilot and define the next steps to move forward with an operational data flow.

Michigan stated while their node was operational; they were unable to add new data flows until their node was upgrade. This upgrade is schedule to be complete by February 22. In the meantime they expressed interest in submitting data via the CDX Test Tool by the end of January. Kristen took an action item to set-up a meeting to discuss this off-line.

Curtis said that with the end of the pilot, he looks forward to the following improvements over the next few years.

1. Implementation of the production system with data publishing services so members of Pacific Northwest Exchange can flow data to the Central Data Warehouse.

2. The next version of the Pacific Northwest Exchange utilizing the OWWQX valid value lists.

3. The ability to get into the backend and production databases, the ability for EPA and other agencies to produce data via Web services as well as the future ability for state nodes to respond directly to Web publishing node requests.

Joe reiterated that the central warehouse is the ultimate goal. The schema provides the ability to look at STORET versus XML submitted data to provide a common view of Water Quality data.

Next Steps

EPA is moving into a phase of reviewing the pilot and pilot schema with the LM Team . Once this review is complete a broader review will be conducted with states. Ideally, some of this discussion will take place at the National Water Quality Monitoring Conference, San Jose, CA scheduled in May.

The Lessons Learned document will be finalized after the February 2-3rd meeting. In addition, a Pilot Handbook is being developed to document the process which the pilot participants went through in order to submit data.

Current Status/Activities

Kristin noted that there should be a few corrections to the minutes:

1. The namespace header issue has not been resolved; the issue transcends data flows. Currently, the guidance on the flows is being standardized in outside efforts.

2. The asynchronous response issue has been resolved since the last meeting.

CDX-TEST TOOL

Ryan provided a brief overview of a document he provide to participants (12/22/05) which describes the CDX interface test tool (the Web interface to CDX) which will allow non-node data submissions.

The steps are as follows:

- o Authentication (log in/password--a default node address will be assigned).
- o Submit the local data file (XML or zip).
- o Check the Status of the Submission (pending or failed).

For the most part information is automatically filled in for the user. If a submission receives a status of failed, users can download a processing report which shows the errors and statistics of the submission.

Draft Pilot Issues List

Participants suggested the following additions to the Pilot issues list:

1. Provide the ability for submitters to query the backend database to assist in deciphering the insert, update, and delete error and warning messages received during submissions.

2. Provide one processing report from CDX, and an automated notification method to the submitter when processing is complete.

3. Enable users to download the processing report or other individual documents without pulling all of the documents from the original submission.

An updated Issue Log will be distributed prior to the next conference call.

The next teleconference will be held at the end of January and will be announced via email.

OWWQX Pilot Biweekly Conference Call
Call Date: March 22, 2006
Time: 12:00 p.m. to 1:00 p.m.
Location: Teleconference hosted by Kristen Gunthardt (EPA)

_____

Joe Wilson (EPA), Tod Dabolt (EPA), Cary McElhinney (EPA), Mike Hart (EPA), Mike Beaulac (Michigan), Jason Smith (Michigan), Kristen Gunthardt (EPA), Dwane Young(EPA), Marybeth Puckace (LMIT), Cathy Anderson (Texas), Julie Lee (Texas), Jeff White (Tetratech/Texas), Mitch West (OR), Curtis Cude (OR), Dan Shoutis(Wind River), Dave Wilcox (Gold Systems), Ryan Jorgensen (Gold Systems), Doug Timms (Enfotech), Ernestine Bryant (LMIT), Mei-Chiun Lee (CSC)

_____

Comments on Pilot Handbook

There was little feedback on the Pilot Handbook. Several of the participants noted that they had not had time to give it a thorough review. Those that had reviewed the handbook found it to be useful.

Comments on Lessons Learned

There was little feedback on the Lessons Learned document.

A question was asked if participants would need training or assistance with Schematron. Dave Wilcox mentioned that this would not be required for most users as CDX will apply the current Schematron rules. Dan from Wind River added that they used Schematron to validate their files locally prior to submitting to CDX and found this to be very useful.

Discussion of WQX Production Schedule & Next Steps

Kristen announced that the monthly STORET call would occur at 12:00pm on Thursday March 23rd. Much of this call would be devoted to the transition from STORET to WQX and Kristen recommended that the pilot participants join in this call.

A 'Future of STORET' link has been added to the EPA STORET website that will hold all WQX Pilot documents including the Pilot Handbook and the Lessons Learned document. To access this site, go to http://www.epa.gov/storet/future_storet.html

There are several upcoming WQX outreach trainings coming up in April. The first of these meetings will be held in Denver on April 18 and 19th   Additional outreach meetings are being scheduled for Chicago, Philadelphia, and Atlanta. A fifth location may be added as well. Pilot participants are encouraged to attend these meeting where possible.

Dan asked Dwane Young if Wind River data submissions to WQX would satisfy the Tribe's requirement to submit their data annually to STORET. Dwane was not sure, as this pilot data was never intended to go into the National STORET Warehouse. Dwane and Dan agreed to look into this issue further outside of the meeting.