APEX DIGITAL SYSTEMS

National Weather Service
The Hydrology XML Consortium

HydroXC
<Hydrology XML Consortium>

Modeling Application-to-Application Data Exchanges using HydroXC XML 2.0

HydroXC Project Phase 2: HydroXC Workshop #1

February 22, 2006

Workshop Agenda

- Welcome and introduction of participants
- Overview: HydroXC activities over the past six months
- Introduction: modeling application-to-application data exchanges using HydroXC XML
- Presentation: Modeling the data exchange between NWS hydrologic forecasting applications FLDWAV and FLDVIEW in HydroXC XML
- Discussion: upcoming HydroXC activities

Table of Contents

Summary

- The National Weather Service (NWS) is currently conducting a longer-term project to develop a common XML schema for transmission of hydrologic data between software applications in the hydrology community. The HydroXC XML schema, draft version 2.0, will be the result of this activity.
- The Hydrologic XML Consortium (HydroXC) is a key part of this work, as it brings together some of the key stakeholders in the hydrology community in the United States, including government agencies, universities and private-sector companies.
- Under the current project, the Consortium will conduct two short workshops to present and discuss the application of the HydroXC schema to existing data transmission problems.
- This document presents material for the first workshop.

Background

- To demonstrate the practical application of the HydroXC schema, version 1.0, the NWS Office of Hydrology (OH) identified two applications that should be retro-fitted to exchange data with each other in XML.
  - FLDWAV: the application generates flood inundation forecast data
  - FLDVIEW: the application generates flood inundation maps based on the data generated by FLDWAV
- The objective of such a design enhancement is to demonstrate how existing applications can benefit from XML integration.
  - Reduce number of input and output files
  - Provide data in open, standards-based format that is generally readably for people and machines
  - Reduce tight coupling of data to science code
- This presentation provides detail on the design of the resulting HydroXC-compliant data exchange interface.

Overview: Existing Data Exchange Format

- FLDWAV writes forecast data to a number of comma-separated text files:
  - One .SCN file contains mapping scenarios that define forecast area over a river system.
  - One .XY file defines each river in the river system, together with cross sections at given river mile points.
  - One .FCS file provides forecast data for each river mile point in each scenario.
- FLDVIEW reads forecast data and generates flood inundation maps for a given scenario:
  - Reads .XY data for a given scenario
  - Reads .FCS data
  - Also reads ENVAR file that provides general system parameters. (These parameters are not included in the proposed XML-based data exchange as they are application-specific).

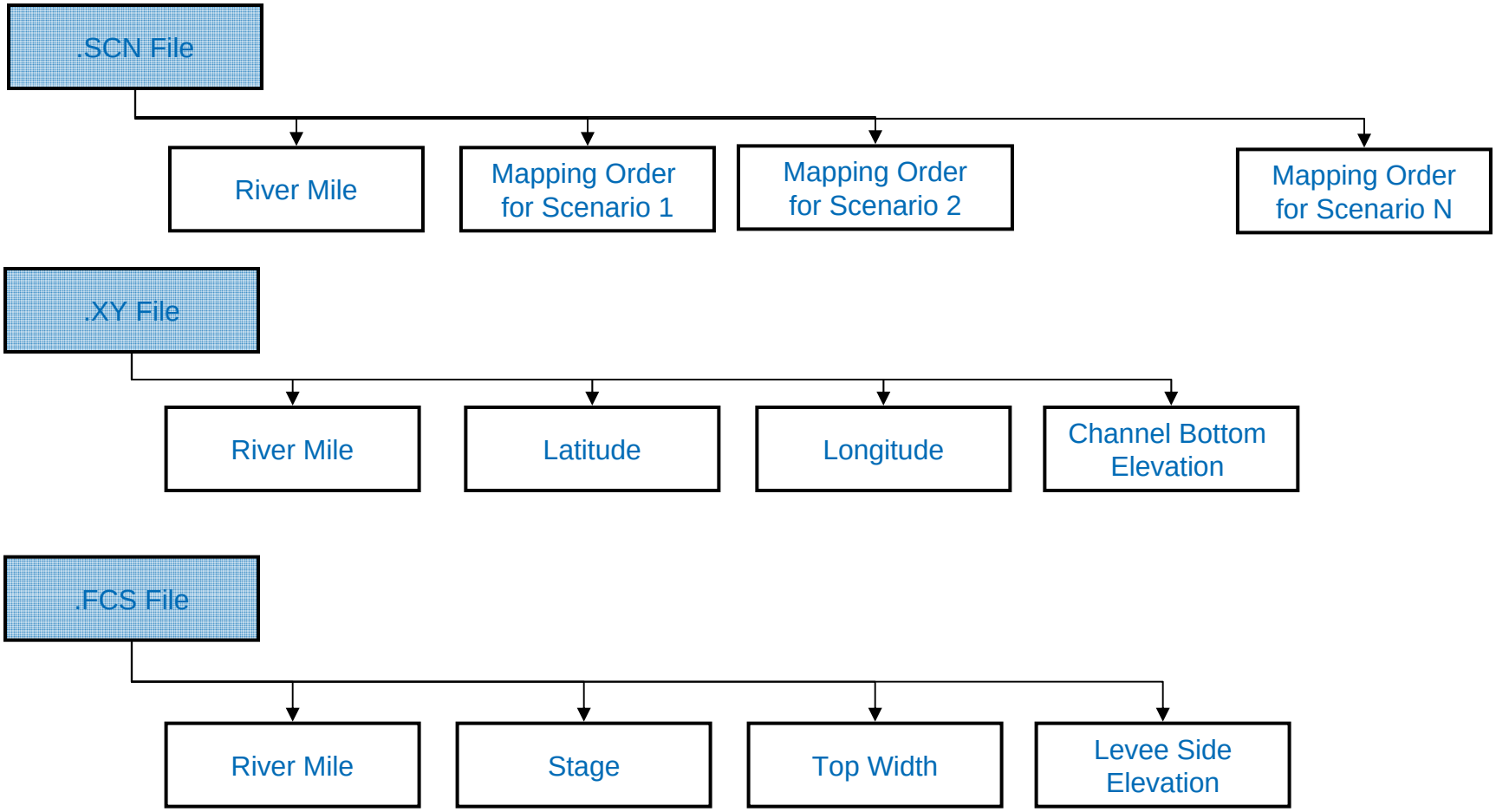
Limitations of the Existing Exchange Format

- The data exchange process implemented between FLDWAV and FLDVIEW is typical for many application-to-application communications that are based on custom data formats:
  - The process requires a number of files, which may reside in different locations. Limitation*:* disconnect of file versions or locations
  - The process requires an external data definition that describes the nature and content of each column in a comma-separated file. Without the data definition, users must resort to code reviews to understand the meaning and use of the data.
  - The data structures (column names, order of columns) are tightly coupled into the consuming science code, and any change in the data structure may have a destructive effect on the consuming code.

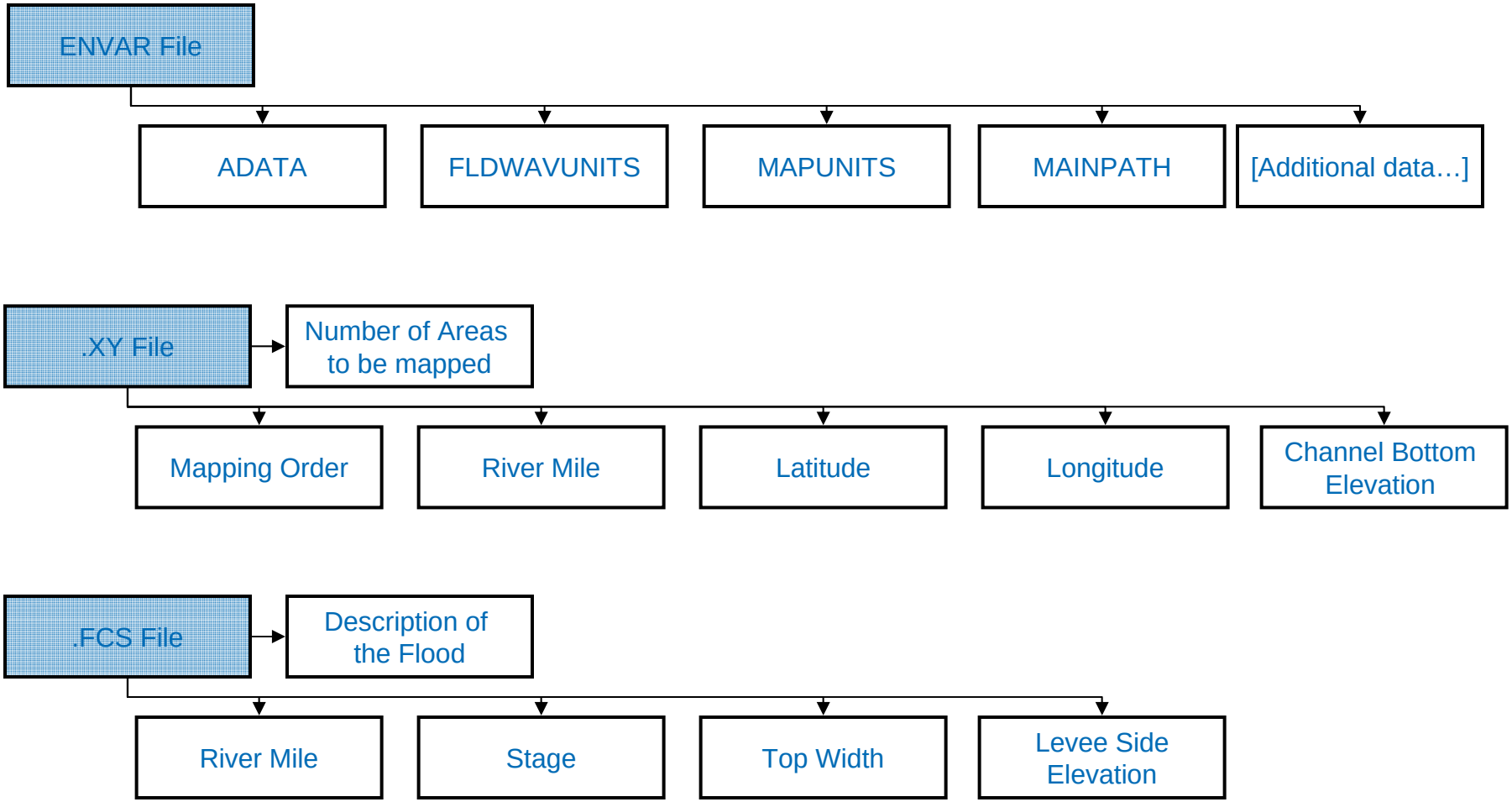
Overview: Proposed Data Exchange Format

- Based on a single HydroXC-compliant XML file that contains
  - Scenario data (.SCN file content)
  - Hydraulic model data (.XY file content)
  - Flood forecast data (.FCS file content)
- Format does not include application metadata
  - ENVAR file content
- The proposed format addresses the existing limitations:
  - It requires only one file, making the process more stable and portable.
  - The file format is self-commenting and self-describing, and includes tags for user and application-level commentary, if needed.
  - The standard schema structure allows loosely-coupled implementation.
- The proposed format can serve as a demonstrator/prototype for other implementations of the HydroXC XML format

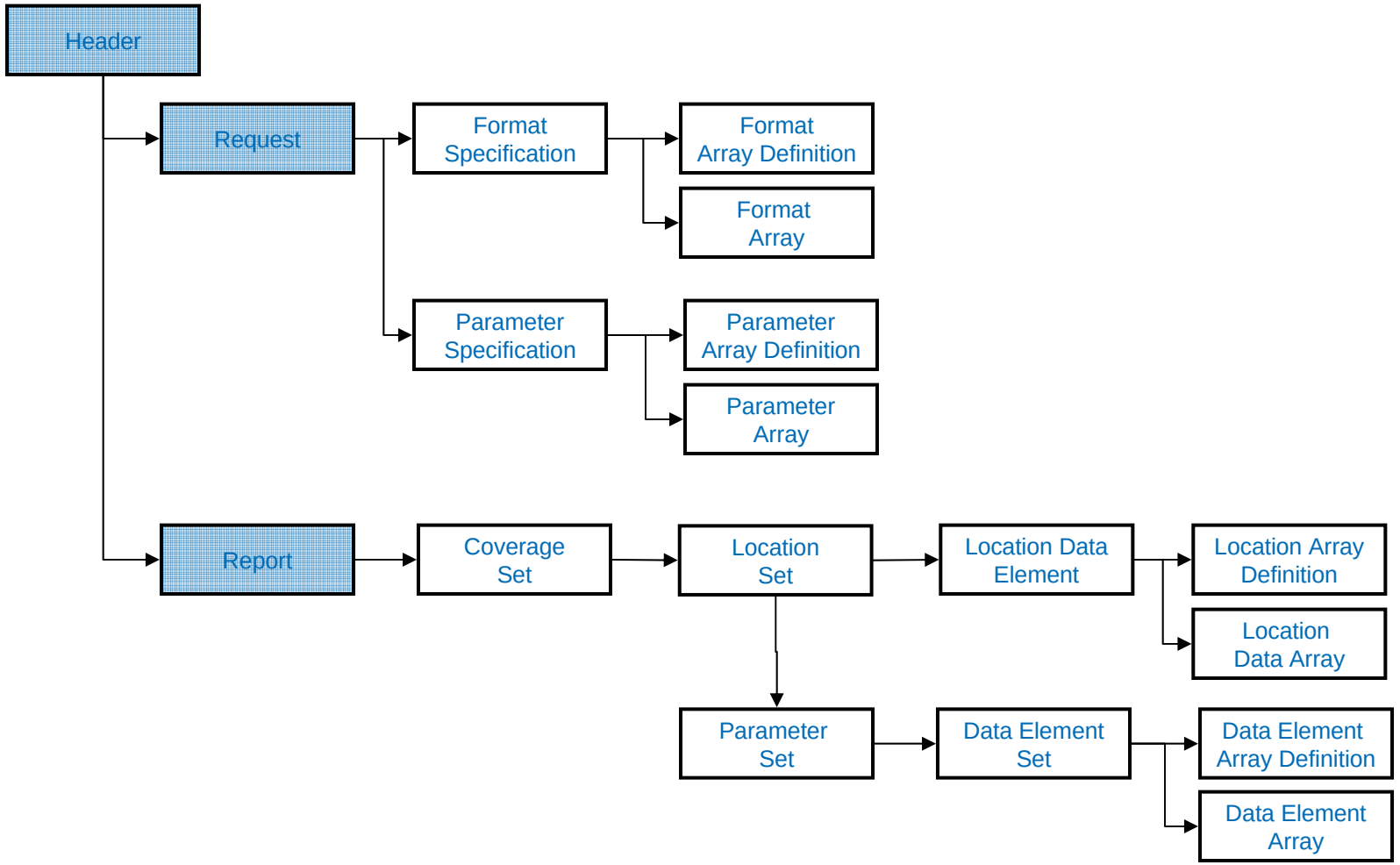
File Structure: Existing FLDWAV Output Files



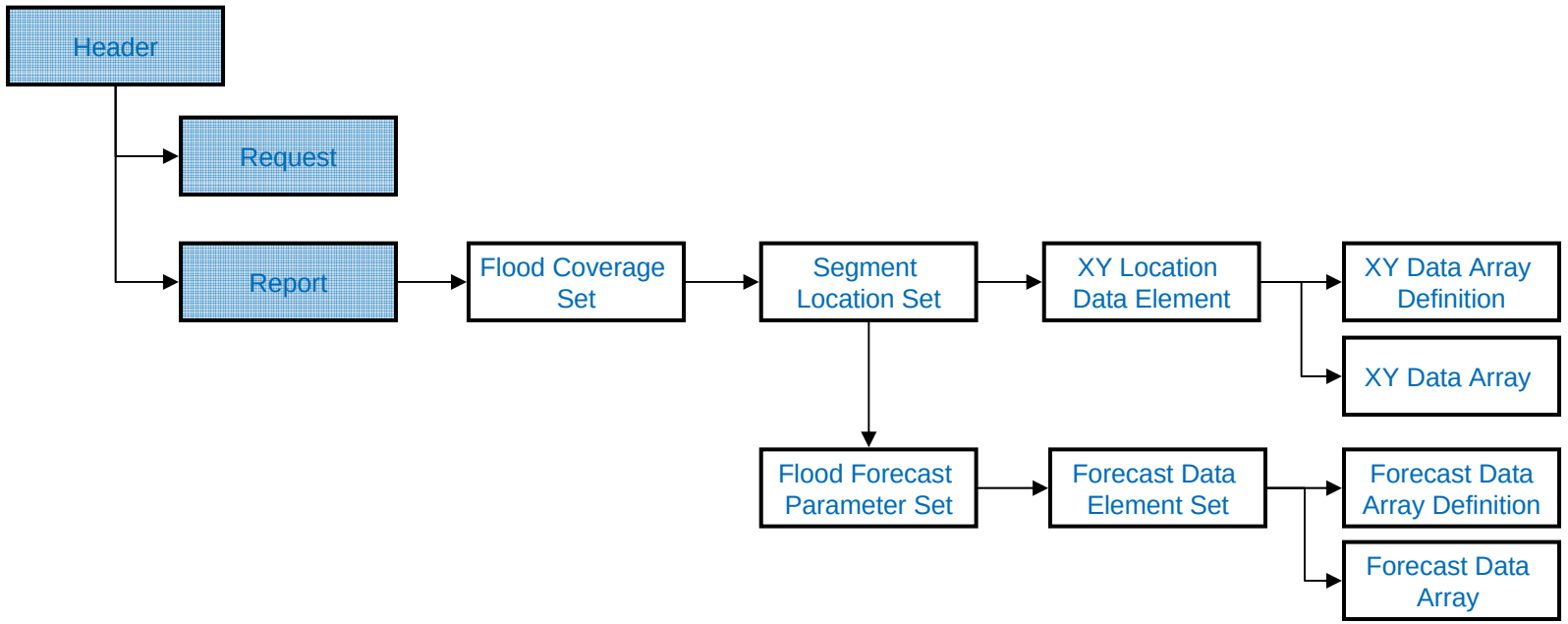
File Structure: Existing FLDVIEW Input Files



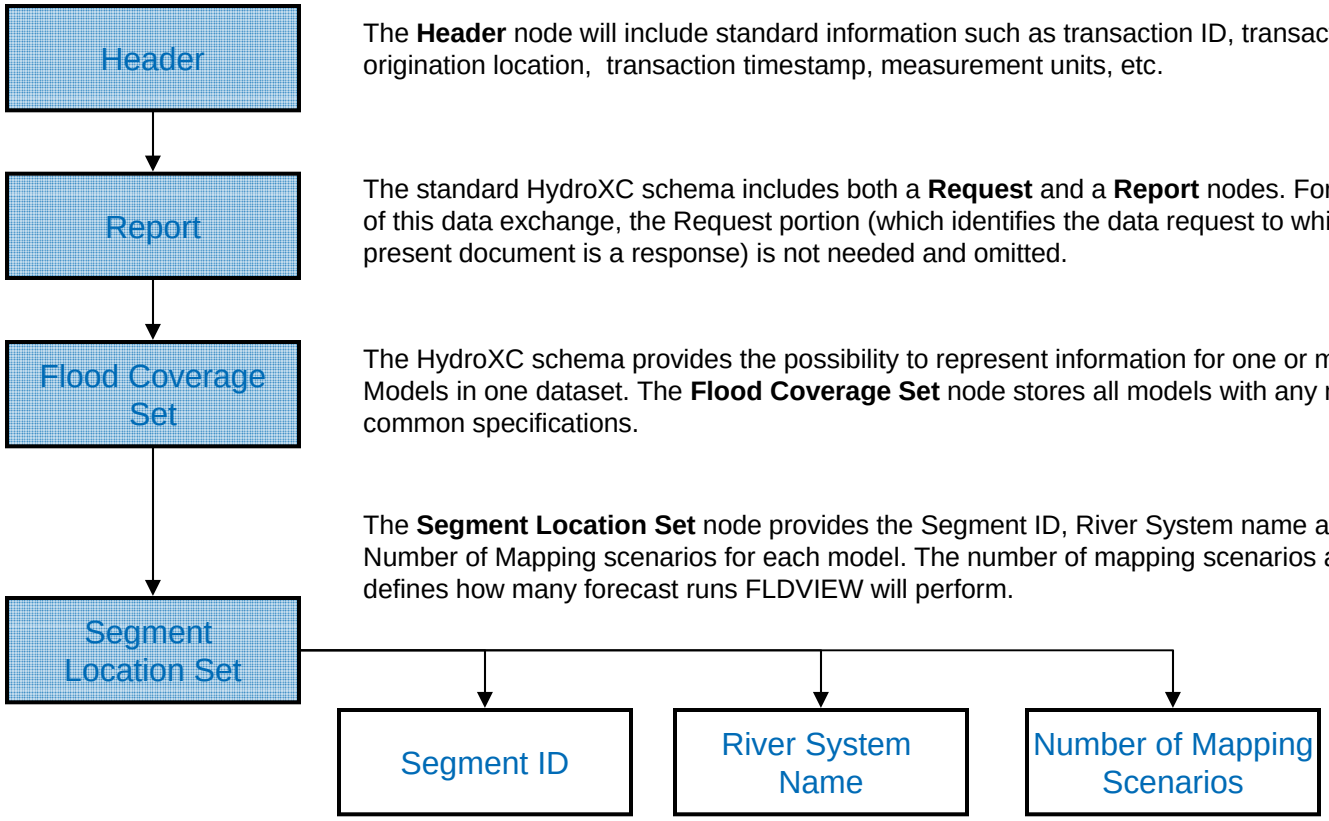
Existing HydroXC XML Schema Version 1.0



Proposed FLDWAV-FLDVIEW Exchange Format



Format Detail: Header, Segment Location Set



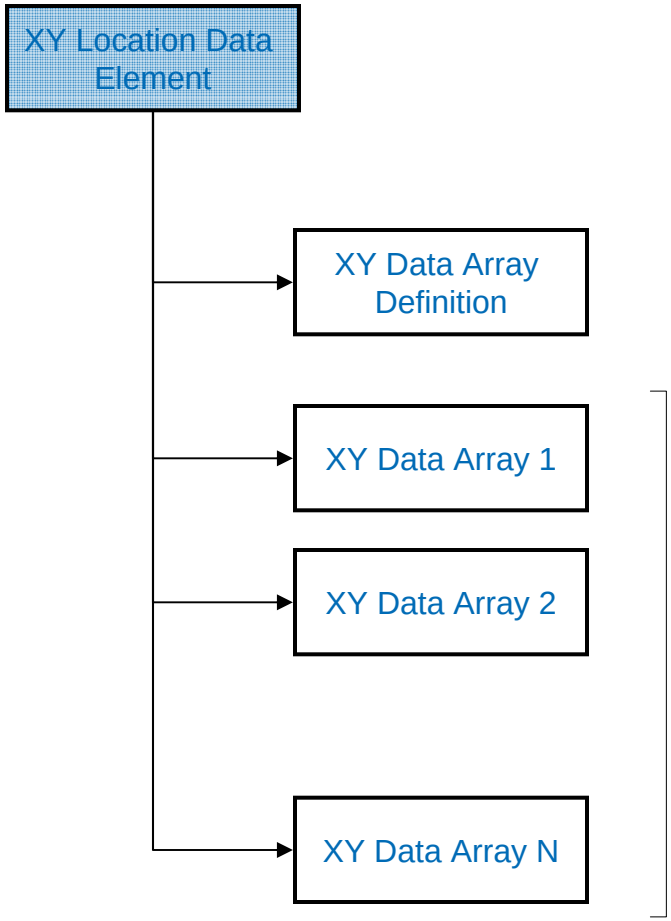
The **Header** node will include standard information such as transaction ID, transaction origination location, transaction timestamp, measurement units, etc.

The standard HydroXC schema includes both a **Request** and a **Report** nodes. For purposes of this data exchange, the Request portion (which identifies the data request to which the present document is a response) is not needed and omitted.

The HydroXC schema provides the possibility to represent information for one or many Flood Models in one dataset. The **Flood Coverage Set** node stores all models with any necessary common specifications.

The **Segment Location Set** node provides the Segment ID, River System name and Number of Mapping scenarios for each model. The number of mapping scenarios also defines how many forecast runs FLDVIEW will perform.

Format Detail: XY Location Element



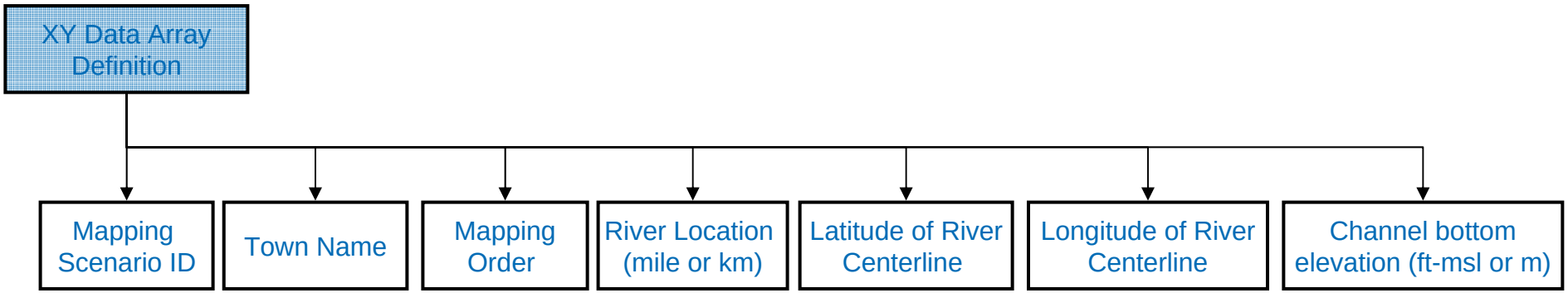
The **XY Data Array Definition** node stores the metadata for the entire collection of data arrays for multiple flood mapping scenarios. For every mapping scenario (represented by a **Data Array**), the node provides a name, description, data type, measurement units and location of the data element.

This collection of one or more **XY Data Arrays** allows management of multiple mapping scenarios for different parts of the river system.

Each Array header includes a numeric Mapping Scenario ID, the town name associated with it and the total number of cross sections covered in the array.

The body section of each Data Array node includes multiple sub-nodes with mapping order and geographical data for every cross section in the river system.

Format Detail: XY Data Array Definition



Mapping Scenario ID is a sequential number mapped area.

Town Name is the name of the nearest town or desired name of the area.

Mapping Order is an order in which rivers will be mapped. 0 - reach will not be mapped, 1 – first river reach to be mapped, 2 - next tributary to be mapped, 3 - next tributary to be mapped, etc.

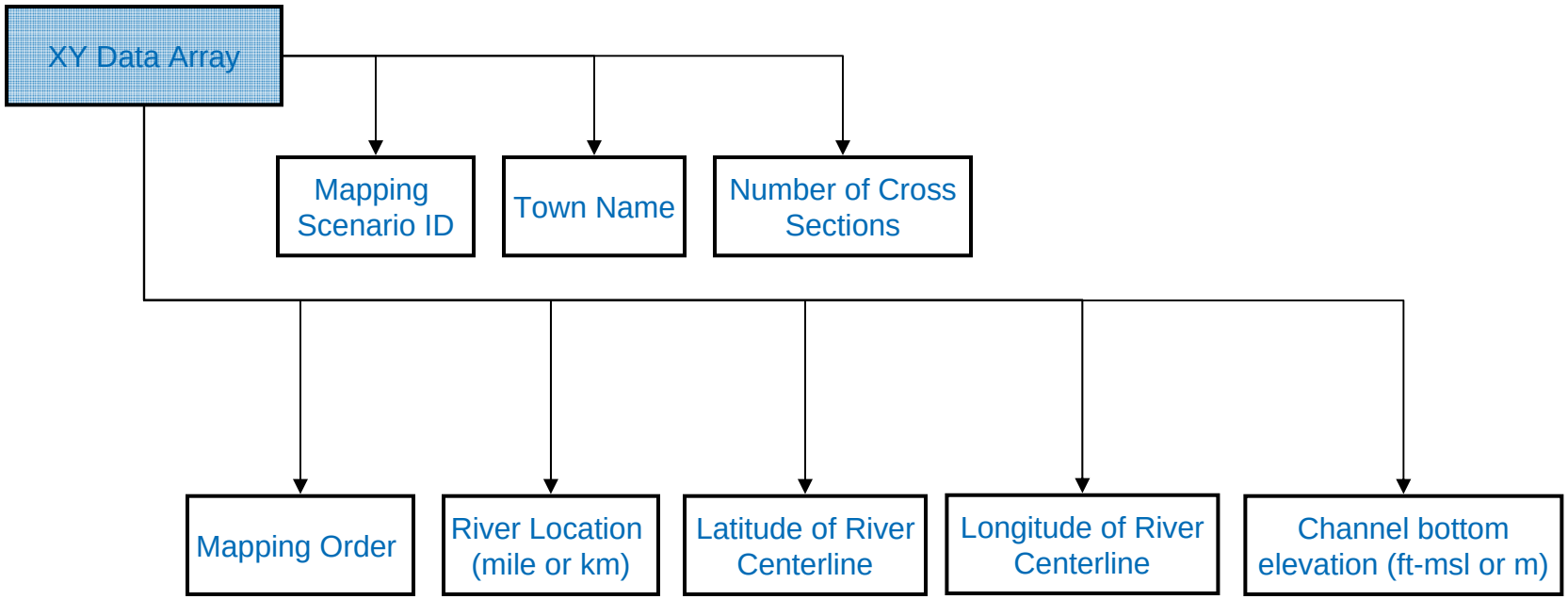
River Location identifies the cross-section river location in miles or kilometers.

Latitude of River Centerline may be zero if not a mapping endpoint.

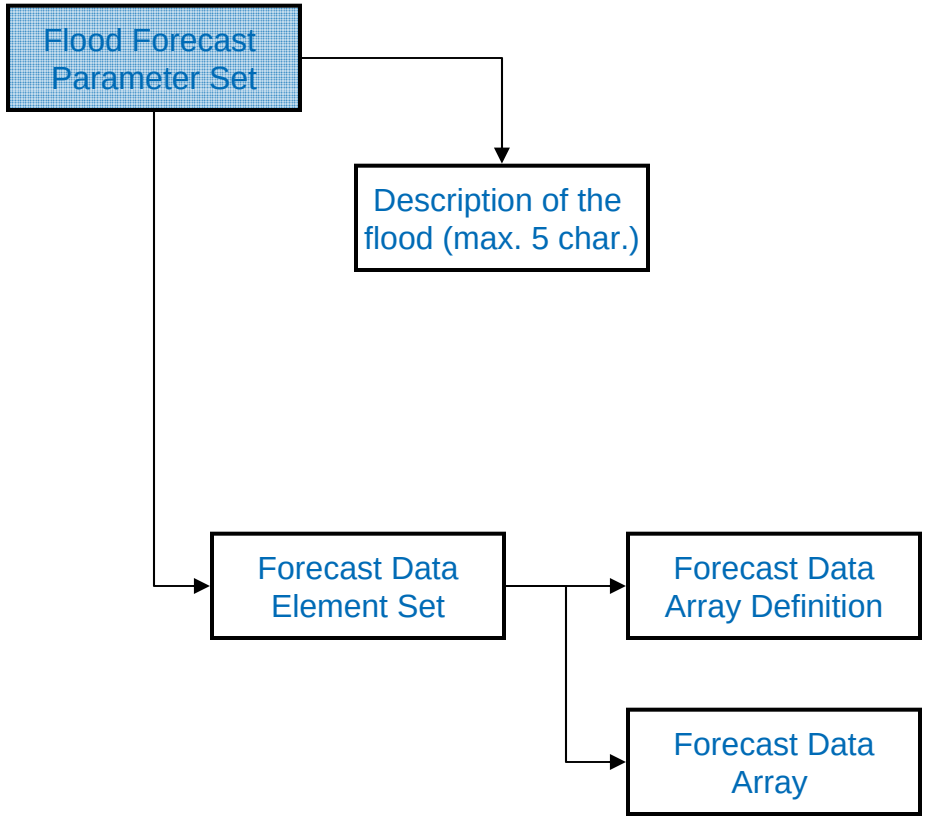
Longitude of River Centerline may be zero if not a mapping endpoint.

Channel bottom elevation (ft-msl or m) up to 4 decimal places.

Format Detail: XY Data Array



Format Detail: Flood Forecast Parameter Set

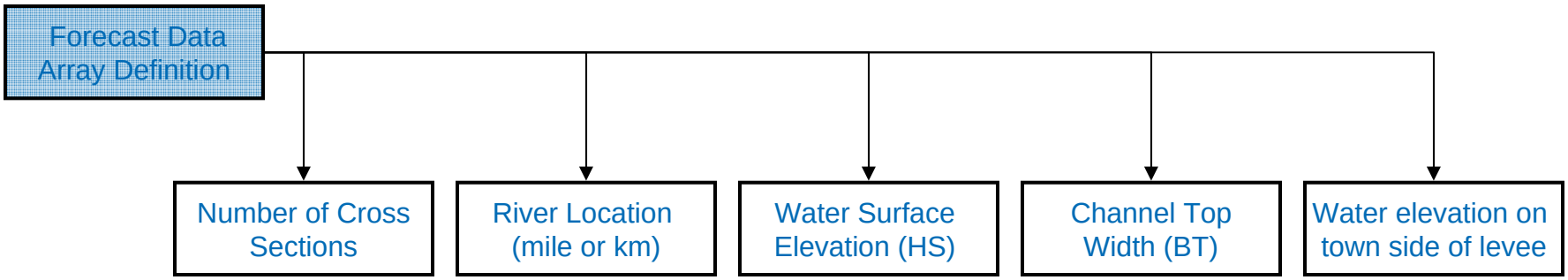


The **Flood Forecast Parameter Set** node specifies both the hydraulic forecast data and its metadata. This structure is able to accept multiple forecasts or time series if needed.

The **Forecast Data Array Definition** node provides descriptions for the metadata for the Forecast Data Array, including the name, description, data type, measurement units and location of the data element within each data array node.

The **Forecast Data Array** contains hydraulic data for each cross section in the river system.
The header of data array node contains the total number of cross sections in the array.
The body of data array node contains multiple sub-nodes with hydraulic data for every cross section in the river system.

Format Detail: Forecast Data Array Definition



Number of Cross Sections defines the number of cross-sections used in the hydraulic flood model.

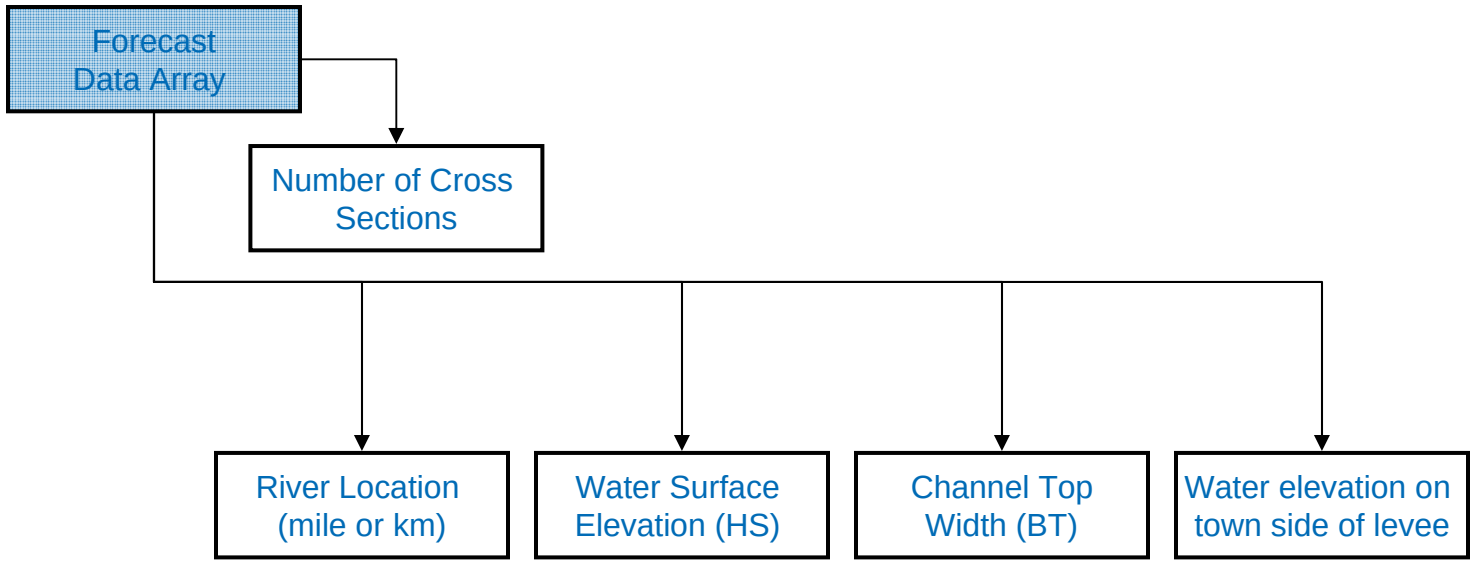
River Location identifies the cross-section river location in miles or kilometers.

Water Surface Elevation is a forecasted elevation in ft-msl or meters.

Channel Top Width represents the channel width corresponding to the water elevation in ft or meters.

If a levee exists, **Water Elevation on Town Side of Levee** will contain the elevation in ft-msl or meters.

Format Detail: Forecast Data Array



Changes from HydroXC Version 1.0 to 2.0

- To support the FLDWAV to FLDVIEW data exchange, we made one modification to the HydroXC Schema:
  - In version 1.0, each location node could contain one data definition node and one data node that provides the data as defined. The data node could only contain one data array. For version 2.0, we are permitting multiple data arrays, each of which provides data subject to the definitions of the data definition node.
  - The same limitation in version 1.0 applied to the node that provides the definition and data for physical elements. We are permitting multiple data arrays under that node as well for version 2.0.

Conclusion & Next Steps

- The design for a HydroXC-based XML data exchange between two operational applications serves as a demonstration of the XML schema's versatility.
- This step is essential in validating the suitability and stability of the schema as an exchange medium. (Note that the schema is entirely message-oriented, i.e., intended for exchange of data between consumers).
- Next Steps should include:
  - Analysis by other Consortium members of suitable applications for XML-based data exchange
  - Implementation of read/write adapters that enable full implementation of data exchange capabilities based on HydroXC XML 2.0
  - Further refinement of the 2.0 schema, to add specificity such as data types, column definitions, display vs. processing names, and other capabilities.