HDDM User's Guide

Richard Jones and R.T. Jones
University of Connecticut
(Dated: February 9, 2010)

The HDDM (Hierarchical Document Data Model) is an xml schema for expressing the meaning and relationships of streaming data from scientific instruments. The design is based on a hierarchical network where each node in the graph has a single parent node, multiple key-value attributes, and an arbitrary number of child nodes, similar to a the elements in an xml document. The model is adapted specifically to the case of repetitive data models such as appear in the data stream from a high-energy physics experiment. The representation of the model in xml is an essential feature, although instantiation in memory does not involve the creation of explicit textual elements or construction of a Document Object Model (DOM) for the data. The HDDM toolkit includes tools to express HDDM streams in xml, check their validity against the schema, and serialize/deserialize from container objects in memory. Originally written in c, HDDM provides application programmer interfaces for C++ and python as well. In addition to its own native data format, applications that use HDDM to access their data can also read/write standard HDF5 files and ROOT trees.