

DFDL WG Call Minutes

This OPEN document will not be filed. It is being kept active.

Meeting about Meetings\OGF

Project	DFDL 1.0
Meeting Date	12-Nov-13 (Tues)
Meeting Time	16:00 - 17:00

Created by Steve Hanson on 09-Mar-11

Last Modified by Steve Hanson on 13-Nov-13

OGF DFDL Working Group Call , 12 November 2013

Agenda

Prepare for your meeting by describing the objectives (both immediate and long-term, if appropriate) of the meeting; and describe key plan details.

1. Daffodil Open Source Project

Status update.

2. Go through public comments

Public comment is now complete.

Continue through the public comments, resolving where possible to do so.

3. Is an optional element in a positional sequence actually a point of uncertainty ?

See email from Steve.

4. AOB

Minutes

Meeting Minutes

Reflect on your meeting as you record all topics and issues discussed , and any tabled conversations . What went well, or what would you do differently next time ? Document those so others can take advantage of your learning .

Attendees

Apologies

IPR Statement

"I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy ."

Minutes

1. Daffodil Open Source Project

Not discussed

2. Go through public comments

No further public comments were discussed

3. Is an optional element in a positional sequence actually a point of uncertainty ?

Question prompted by noting IBM DFDL behaviour when such an element is present but contains data that causes a processing error (eh, type is xs:int and data is 'abc'). The IBM DFDL parser backtracks and tries the next alternative, which is to look for the next separator, which it does not find and so causes another processing error. The DFDL spec says that this is a potential PoU and is also an actual PoU. Is that correct? It was agreed that this particular case should not be made an exception, and it remains an actual PoU. However implementations are free to optimise this situation in order to output better diagnostics etc.

Meeting closed

Next regular call

Tues 19th Nov 16:00 UK

Create Action Items

Record the to-do's and individuals assigned by entering the appropriate information in the form below. Press the "Create Action Items" button to create specific to do's that can be tracked in the assignee's Work for Me views. " All Action Items will be tracked in the Action Items and Other Meeting Documents tab.

Action Items and Other Meeting Documents

Subject	Document Type	Created	Modified

Next action: **242**

Actions raised at this meeting

--	--

No	Action

Current Actions :

No	Action
066	<p>Investigate format for defining test cases (All)</p> <p>25/11: IBM to see if it is possible to publish its test case format .</p> <p>04/12: no update</p> <p>...</p> <p>17/02: IBM is willing in principle to publish the test case format and some of the test cases . May need some time to build a 'compliance suite'?</p> <p>24/03: No progress</p> <p>03/03: Discussions have been taking place on the subset of tests that will be provided .</p> <p>10/03: work is progressing</p> <p>17/03: work is progressing</p> <p>31/03: work is progressing</p> <p>14/04: And XML test case format has been defined and is being tested.</p> <p>21/04: Schema for TDML defined. Need to define how this and the test cases will be made public</p> <p>05/05: Work still progressing</p> <p>12/05: Work still progressing</p> <p>02/06: Work still progressing on technical and legal considerations</p> <p>...</p> <p>25/08: Will chase to allow Daffodil access to test cases . The WG should define how implementation confirm that they 'conform to DFDL v1'</p> <p>01/09: IBM still progressing the legal aspect. Intends to publish 100 or so tests as soon as it can, ahead of a full compliance suite.</p> <p>08/09: IBM still progressing</p> <p>15/09: IBM still progressing, expect tests to be available within a few weeks</p> <p>22/09: IBM still progressing, expect tests to be available within a few weeks</p> <p>29/09: Test cases are being prepared.</p> <p>06/10: Some test cases should be available next week. Steve would like to be able to show the test case information at OGF 30.</p> <p>13/10: Still progressing</p> <p>10/11: Legal issues cleared, IBM in process of collecting 100 example test cases, ideally ones that fit the 'extended conformance' of NCSA Daffodil</p> <p>17/11: Work is progressing on verifying the test cases. It should be possible to distribute to the WG in 2 weeks.</p> <p>24/11: About half the test cases have been completed and are being reviewed internally.</p> <p>01/12: Test cases should be available shortly</p> <p>08/12: The test cases are in internal IBM review. Probably need a bit of reorganising before publication</p> <p>Stephanie gave a brief overview of the format of the test cases.</p> <p>15/12: Ruth joined the call to provide the latest status. The test cases have been updated and a draft read.me produced. Although not ready for public distribution Ruth will send them to Joe for feedback.</p> <p>22/12: Test cases were sent to Joe for initial testing which found some problems in the Daffodil parser</p> <p>12/01: All current tests use a default format whih Daffodil doesn't currently support. Joe suggested that there should be test that defined the same function using different definition forms. Also suggested that default formats should be provided by the WG. This had always been the intention. Action 133 raised to track.</p> <p>19/01: There is currently no resource available in IBM to make more tests available . IBM to discuss how/if it can make a 'minimal compliance test suite' available.</p>

	<p>26/01: Action kicked off within IBM. There was a brief discussion about naming and organisation of test cases but no preferences were expressed</p> <p>02/02: IBM will not have the resources to develop a full test suite in the near future. Steve suggested that we produce a list of required test cases so that anyone could supply them.</p> <p>09/02: Steve had previously sent a list of areas to be tested. Please review.</p> <p>23/02: Please review Steve's list of areas to be tested</p> <p>02/03: Alan had reviewed Steve's list and we went through his comments. Agreed there is no need for separate tests for the info set or for dfdl: property lists, unions etc but comment will be added that these should be exercised during property testing.</p> <p>09/03: Alan updated the test document. Need more introduction and perhaps adopting the OGF template.</p> <p>30/03: Ownership of test document passed to Steve. This action is merged with 112 and will cover all aspects of compliance suite.</p> <p>13/04: IBM will not have time to create a compliance suite in the near future. Probably best to make this action deferred for now.</p> <p>...</p> <p>10/07/2012: Discussed schemes to create interchangeable tests. Ideally need a DFDL defined error code per failure, in conjunction with specific inserts.</p> <p>...</p> <p>26/3/2013: Resurrecting deferred action.</p> <p>We have got to the point where it makes sense to converge the IBM DFDL and Daffodil variations of .tdml file.</p> <p>Steve to seek permission from IBM to make the list of IBM DFDL error messages available to DFDL WG.</p> <p>...</p> <p>24/5: No further progress.</p> <p>28/5: Mike summarised the status of Daffodil's tdml runner. Since IBM shared the tdml format, Daffodil has added a) bit file support with in-line comments; b) embedded schema; c) failure checking by multiple string matching. IBM has added a) some flags that map to parser API 'features' such as optional checks; b) code to handle illegal XML characters. 1200 parser test cases written for Daffodil, about 60 of the original IBM shared tests now pass in Daffodil. Steve will email OGF and ask if there is an approved process for demonstrating that multiple implementations generate the same set of test results. To progress with a shared tdml format, IBM will need to get legal approval to view the Daffodil source test cases, Steve to kick this off. Mark noted that IBM's tdml format has evolved in order to make the info set comparison easier, Mark will see whether the shared tests use the latest version.</p> <p>4/6: Steve has emailed OGF for guidance, reply received. Experience documents needed to verify conformance, but there is not a requirement to have executable tests. However, a set of executable tests is what we need ideally.</p> <p>Discussed error messages and identifiers for different errors and what the granularity should be. Steve has asked for permission to send the IBM DFDL error messages to the DFDL WG, they should be used as a starting point. Need to agree what constitutes the minimum content of an error message.</p> <p>...</p> <p>22/10: No further progress</p> <p>31/10: Steve has permission to send IBM DFDL's error messages to the WG.</p> <p>5/11: Error messages sent</p> <p>12/11: No further progress</p>
200	<p>Establish recommended practices for pushing changes to GitHub (Mike)</p> <p>29/1: Mike will talk to Tresys who have used Git a lot.</p> <p>5/2: Mike to talk to Tresys this week, Tim has sent some links.</p> <p>12/2: Information sent by Mike, Steve to review.</p> <p>...</p> <p>12/11: No further progress</p>
224	<p>Add section for implementation defined limits (Jonathan)</p> <p>3/9: Several places in the spec cite this, should be grouped. Currently partially listed in section</p>

	<p>2.6. Also note distinction between 'implementation defined' and 'implementation dependent'. Check spec for correct usage. Resolve during public comment. 10/9: No progress 17/9: Jonathan sent a reference to the W3C XProc standard where the distinction is made clear. Jonathan will go through the spec and gather everything that is implementation defined/dependent. Public comment to be raised 24/9: With Jonathan to raise. 1/10: Public comment 97 raised (http://redmine.ogf.org/boards/15/topics/97) 8/10: With Jonathan to provide words. 22/10: Jonathan has defined implementation defined/dependent and started to classify. Steve and Mike had trouble with the definitions, Steve to re-word and send for comment. 31/10: Reworded version sent 5/11: Rewording approved. Jonathan proceeding with classification, will distribute for review when complete. 12/11: Still with Jonathan</p>
228	<p>Review set of tutorial lessons (All) 17/9: Lesson 1 proposes a set of lessons, needs reviewing as over 2 years old. ... 22/10: No progress 31/10: Becoming a focus for Tresys. Steve to send his 'Modeling Data Formats using DFDL' powerpoint. ... 12/11: No further progress</p>
237	<p>Public comment : Do we need the lengthKind 'endOfParent' restriction for local groups ? (All) 5/11: http://redmine.ogf.org/boards/15/topics/51. Stops such an element being wrapped in a local group. It looks like this was added to make checking easier. Really the restriction is that the element is "last in the box", and could in principle have surrounding intermediate sequences or implicit choices. Need concise language to express this. 12/11: Need language written.</p>
239	<p>Public comment : Implicit alignment and groups (All) 7/11: http://redmine.ogf.org/boards/15/topics/45 and http://redmine.ogf.org/boards/15/topics/57. Discussed whether implicit alignment of a model group should always be '1', to avoid problems of having to look ahead to child elements, which is problematic if the group has an initiator or there are alternatives such as choice branches or unordered content. Led onto discussion on the existing rule that an optional element must have same alignment as the following element (define 'following'). Should there be a rule for sequences that are unordered or have floating children must have children with same alignment? This is consistent. Noted that these rules are all meant to stop authors from creating silly schemas. Alternatively should such checks be relaxed? 12/11: Email sent to Suman, Steve to chase.</p>

Closed actions

No	Action
235	<p>Public comment : Request clarification of escape schemes (Mike) 22/10: http://redmine.ogf.org/boards/15/topics/42. Started on this but did not finish, Mike to send notes by email. 31/10: Notes received, review for next call 5/11: No further progress 12/11: Closed. Mike's notes augmented by Steve via email. Talked through the email, agreed with the answers to the public comment questions, further updates by Mike. Public comment updated. Resolved.</p>
236	<p>Public comment : Test ICU behaviour for unpaired surrogate code points (All)</p>

	<p>31/10: Public comment http://redmine.ogf.org/boards/15/topics/49. Is ICU able to implement the specified behaviour?</p> <p>5/11: No further progress</p> <p>12/11: Closed. Tests show that ICU can be configured to perform the specified behaviour. Resolved.</p>
238	<p>Public comment : Number pad character and alternative '0' characters (All)</p> <p>5/11: http://redmine.ogf.org/boards/15/topics/54. Determine the set of zero-equivalent characters and consider whether we can pad/trim with respect to the set, or specify any one of them, or other possible ways to resolve this.</p> <p>12/11: Closed. To solve this properly requires DFDL to provide a 'numbering system' property, akin to <code>dfdl:calendarLanguage</code>, which would apply to Number and Calendar types. Decided not to provide such a property in DFDL 1.0. Resolved.</p>
240	<p>Public comment : Define collation units (All)</p> <p>7/11: http://redmine.ogf.org/boards/15/topics/60. Agreed that DFDL spec needs to say something about collation, if just to say that DFDL string comparisons only ever use a default collation. This makes sense as otherwise schemas become non-portable. Suggests that we should drop the XPath functions that take a collation as an argument.</p> <p>12/11: Closed. Noted that XSDL 1.0 does not mention collation anywhere. DFDL spec can explain collation units in terms of characters, and can state that default Unicode collation algorithm is used. Functions with explicit collation arguments will be dropped. Resolved.</p>

Deferred actions

No	Action
131	<p>Transformation of DFDL properties to a canonical form (Joe)</p> <p>08/12: Joe has produced a XSLT to transform a DFDL schema to a canonical element form. When tested it should be made available on the WG gidforge site.</p> <p>15/12: Alan tested against test dfdl schema which worked correctly (after fixing some errors in the schema)</p> <p>22/12: no update</p> <p>12/01: Joe has some defects to fix before making available on gridforge.</p> <p>19/01: There is a difficult problem to solve before Joe make the style sheet public</p> <p>26/01: Working on problems</p> <p>02/02: no progress</p> <p>09/02: As it wasn't a simple as exoected this will be treated as a low priority action</p> <p>23/02: Low prioity</p> <p>09/03: Low priority</p> <p>30/03: Deferring for now</p>
233	<p>Public comment : Formats with bit order reversed (Mike)</p> <p>1/10: http://redmine.ogf.org/boards/15/topics/43. Mike to provide words for potential new property for review.</p> <p>8/10: Words sent by Mike generated considerable discussion. Mike will update the words to make the subject more consumable, and move the bulk of the discussion to a new main section at the end of the spec (suggest between existing sections 24 & 25).</p> <p>22/10: Mike wants to have a working implementation before closing on this, so marking the public comment as deferred.</p> <p>31/10: Deferring for now</p>
241	<p>Public comment : Bi-di properties placement in precedence section (All)</p> <p>7/11: This looks deliberate but the asymmetry between parsing and unparsing is unclear. Really needs Daffodil or IBM DFDL to implement these properties, which has not happened yet. Deferring this action.</p>

Work items:

No	Item	Owner	Target	Status

045	Resolve public comments and incorporate into spec	All	2013-10-22	Pending
-----	---	-----	------------	---------