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	22-Sep-15 (Tues)
Meeting Time	16:00 - 17:00

Created by Steve Hanson on 09-Mar-11
Last Modified by Steve Hanson on 22-Sep-15

OGF DFDL Working Group Call , 22 September 2015

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. 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
Mike Beckerle
Steve Hanson
Andrew Edwards
Alex Wood

Apologies

Minutes

1. Daffodil Open Source Project
Work continuing.

2. DFDL schemas for X 12
IBM now able to generate these from the X12 specification tables.
Similar structure to the DFDL schemas for EDIFACT.
Now includes message-level error recovery to allow entire Interchange to be parsed.
Mike to see if Daffodil supports these schemas.

IPR Statement
“I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy .”

Meeting closed
16:40 UK

Next regular call
Tues 6th October 2015 @ 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: **285**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
228	**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. ... 19/11: No further progress 26/11: Possibility of help from MITRE high-school student, and from Marisa at IBM. ... 11/3: No further progress 25/3: MITRE have produced a couple of new tutorials under the guidance of James Gariss. Jonathan to forward for review. Mike observed that an html tutorial could be generated from a tdml file using XSLT. 11/4: Not discussed 15/4: Jonathan will send 4 new mini-tutorials. Need to figure out best way to incorporate into the tutorial structure. 29/4: Tutorials received. Mark has taken a quick read. Mark & Steve to review and report back. 6/5: Still with Mark and Steve 20/5: Mark has reviewed. Will ask IBM information development to recommend a way to portray the existing and new lessons, preferably web-based. Find somewhere to host them. OGF? GitHub? developerWorks? NCSA? 3/6: Steve has also reviewed. ... 17/6: No further progress on tutorials. Tim is looking into the creation of some DFDL how-to videos using the IBM Integration Studio. ... 31/3: No further progress 14/4: Agreed that the need for better tutorials has become pressing for Daffodil users who aren't using IBM's tools and material. Discussed creating tutorials based on a tdml file with comments that is processed to produce html. Mike to investigate. 28/4: Mike has sent an example tdml file which embeds instances of a new 'tutorial' element in various places. These elements contain html which can be extracted and formatted in a browser. Suggest future DFDL tutorials are created using this technology. 12/5: Not discussed ... 22/9: No further progress
250	**Standardise on a single tdml format for DFDL tests (All)** 5/2: Steve has requested permission for IBM to view / use the Daffodil tdml files, as a precursor to trying to standardise on a common tdml format. Was formerly part of action 066. ... 18/2: No further progress 11/3: Mike and Steve discussing the best way to share and cooperate on tdml format. 25/3: Discussed the creation of an OGF document that will own and define a standardised tdml format. 11/4: Proposal is for the OGF document to define a tdml format without Tresys or IBM copyright statement. 15/4: Draft document on Redmine ... 6/5: No further progress 20/5: Mark has read through the document. Particularly concerned with how namespaces are handled in the infoset. ... 17/6: No further progress 25/6: Mike has added bit order capability as per action 233. ... 9/12: No further progress 6/1/15: Mike to resurrect this as Tresys would like to run their tdml suite against both Daffodil

	and IBM DFDL. ... 10/2: No further progress 24/2: Mike updating the Daffodil TDML test runner to handle unparser (ie, serializer) tests ... 14/4: No further progress 28/4: Tresys have enhanced their tdml runner to allow unparser tests and round-trip tests (parser->unparser->parser) as well as the new tutorial tag (see action 228) 12/5: Not discussed ... 22/9: No further progress
279	**Improve defaulting description to explicitly cover local groups (Mike)** 28/4: Only talks about elements, should mention local sequence and choice. 12/5: Not discussed 23/6: Section 15.1.3 needs to say what happens when a choice branch does not contain any elements; such a choice branch is selected (but see action 280 below as minOccurs '0' might change this). Section 9.4 also needs updating to say what happens when local groups are found within a complex type. 11/8: Steve did some tests with IBM DFDL. Just need some words as above. Action assigned to Mike. 25/8: In progress 22/9: No progress
282	**Does XPath have operators for checking a value is in a range ? (Steve)** 12/5: Investigate whether equivalent to DFDL4S 'in' operator exists. 23/6: Mike has found an XPath 'intersect' operator. Handles the enumeration case well, but not as convenient for ranges as DFDL4S's 'in' operator. 11/8: Looked back at the motivating example from DFDL4S. Agreed that DFDL functions to do the equivalent of 'in' and 'inrange' would be useful if nothing can be re-used from XPath. Steve to write up a proposal taking into account different data types. .. 22/9: Proposal sent by Steve for new functions dfdl:checkValues(), dfdl:checkRangeInclusive(), dfdl:checkRangeExclusive(). Discussed whether both range functions needed, and whether they should be allowed for float and double. Mike noted that dfdl:checkConstraints() could be used instead of all three functions if values were static. Follow-up with DFDL4S team to see if they had thought of that.

Closed actions

No	Action
280	**Choice branch rule minOccurs <> 0 ? (All)** 28/4: Is this still needed given how occursCountKind and discriminators now work ? 12/5: Not discussed 23/6: Steve sent an email where he attempted to find the justifications for the restriction (extra point of uncertainty, known-to-exist rules). Steve will do some more digging to see at what point this restriction was introduced, and look at XML scenarios. 11/8: Steve believes that the restriction is to be consistent with not allowing minOccurs =0 on the choice itself, as both allow the choice to be absent logically. Think about this more for next call. Also is the inputValueCalc restriction needed? 25/8: Steve has emailed his thoughts. Mike in agreement. Steve to check with Alex who raised the problem originally. 22/9: **Closed.** Alex also agrees with the email. Clarification to be added to choice section to describe behaviour when dfdl:occursCountKind is 'expression' and 'parsed'. **Erratum raised, tracked by https://redmine.ogf.org/issues/299.** Property dfdl:inputValueCalc to be allowed on root branch of a choice. **Erratum raised, tracked by https://redmine.ogf.org/issues/300.**
283	**Investigate incorporation of ICU fallback mappings into dfdl :encodingErrorPolicy (Steve)** 23/6: IBM has requirements from customers to support fallback mappings.

	11/8: No progress. Steve to look at prior proposal. 25/8: Discussion and proposal emailed by Steve. Should DFDL add a new property or can it all be handled by extra dfdl:encodingErrorPolicy enums? 22/9: **Closed.** Proposal sent by Steve after discussion within IBM DFDL team. Will adopt proposal A so two new enums "fallbackOrError" and "fallbackOrReplace" will be added to dfdl:encodingErrorPolicy. **Erratum raised, tracked by https://redmine.ogf.org/issues/301.**

Deferred actions

No	Action
131	**Transformation of DFDL properties to a canonical form (Joe)** 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. 15/12: Alan tested against test dfdl schema which worked correctly (after fixing some errors in the schema) 22/12: no update 12/01: Joe has some defects to fix before making available on gridforge. 19/01: There is a difficult problem to solve before Joe make the style sheet public 26/01: Working on problems 02/02: no progress 09/02: As it wasn't a simple as exoected this will be treated as a low priority action 23/02: Low prioity 09/03: Low priority 30/03: Deferring for now
200	**Establish recommended practices for pushing changes to GitHub (Mike)** 29/1: Mike will talk to Tresys who have used Git a lot. 5/2: Mike to talk to Tresys this week, Tim has sent some links. 12/2: Information sent by Mike, Steve to review. ... 2/12: No further progress 14/1: Deferring until needed
241	**Public comment: Bi-di properties placement in precedence section (All)** 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. ... 23/9: Candidate to be moved out to 1.1 ?
242	**Public comment: dfdl:valueLength and dfdl :contentLength descriptions (Mike)** 19/11: http://redmine.ogf.org/boards/15/topics/63. Agreed that the function names were ok as per errata 3.18, and that the spec is clear that they refer to the grammar regions. However the grammar regions mentioned do not fully include literal nil values. Discussed what happens when parsing - remember the length or re-parse? What about lengthUnits 'characters' when the data is binary? Also the 'Notes' that follow the table need to be reworked. 26/11: Needs wording to handle all the issues found, assigned to Mike. ... 11/3: Still with Mike 25/3: Mike has sent out revised wording, reviewed by Mark and Steve. Noted that the words need to explain the concept of building a complex element from the bottom up, and these words are equally applicable to several places in section 12.3. Mike to revise accordingly. 11/4: More revised wording sent by Mike. Started to review but realised it needed some off-line preparation and thought. Review for next call. 15/4: Review comments from Steve and Tim. The functions need to be clear that they work off the infoset value. The detailed wording is needed but should be removed to a new sub-section of 12.3, probably at end. Most sub-sections of 12.3, and the functions in 23.5.3 will refer to this new sub-section. 23.5.3 should limit itself to behaviour specific to the functions, such as not

	potentially represented, the effect of the $lengthUnits argument. Also discussed what happens if $path argument returns a nodeset > 1; should be a processing error, can always use a predicate to select one node of an array. 29/4: See various email discussions. Several things noted by Mike, and he recommends a rewrite of some of section 12.3. Then the description of the two functions becomes much simpler. Deferring for now, and will resurrect after current spec revision is finalised. 6/5: Mike is working on a mind map for the length section. Deferring until needed. .... 23/9: Rewrite should be postponed to future 1.1. Still need to answer the original questions about the functions though...
251	**Create official error codes (All)** 5/2/14: Create official error codes for all possible errors implied by the DFDL spec. This is a big piece of work, so this action is deferred for now. Was formerly part of action 066. ... 10/2/15: Un-deferring this action for consideration. The IBM error message numbers could be a good starting point for a list of error codes, although they are a mixture of generic and specific so some revision likely to be needed. Tresys to evaluate. ... 24/2: No progress 24/3: Mike has read through the IBM error messages and will write up findings 31/3: No further progress 14/4: Mike has sent email. Concluded that this is a large exercise and effort better put in to tutorials. Deferring again.

Work items:

No	Item	Owner	Target	Status