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	25-Jun-14 (Wed)
Meeting Time	16:00 - 17:00

Created by Steve Hanson on 09-Mar-11
Last Modified by Steve Hanson on 26-Jun-14

OGF DFDL Working Group Call , 25 June 2014

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. US MIL-2045
Mike has created a DFDL experience document that covers the modelling of this family of standards:
http://redmine.ogf.org/dmsf_files/13268.
It includes potential enhancements to DFDL that are needed in order to do this correctly and/or efficiently.

3. 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
Steve Hanson
Mike Beckerle
Andy Edwards

Apologies
Tim Kimber
Jonathan Cranford

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. US MIL-2045
Mike's DFDL experience document for this proposes several additions to DFDL. Each will be considered in turn. Some are needed for DFDL to be able to handle MIL-2045 formats, others make the modelling easier. First step is to review the document and add comments before the next WG call in 2 weeks time. **Action 233 un-deferred.**

3. Direct Dispatch choice - case insensitive key comparison
Andy pointed out that making this a case insensitive comparison is more convenient for the user but affects the performance of the dispatch, and this is intended to be a performance sensitive feature. A user with mixed case keys can always use fn:toLower() in the expression.
Noted that the dfdl:ignoreCase property needs to borne in mind when resolving this.
Andy will send a proposal to the WG mailing list.

Meeting closed
17:15 UK

Next regular call
Tues 8th July 2014 @ 16:00 UK
Next week's call cancelled

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: **263**

Actions raised at this meeting

No	Action

Current Actions :

No	Action
224	**Add section for implementation defined limits (Jonathan)** 3/9: Several places in the spec cite this, should be grouped. Currently partially listed in section 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. ... 28/1: Still with Jonathan 5/2: Jonathan is up to section 12.7. Discovered an issue with binary packed calendars, new **action 252** raised. ... 11/2; No more progress 18/2: Jonathan has around 20 changes identified so far, and has sent for an initial review. Comments back to Jonathan before next week's call please. 11/3: Reviewed the document so far. Decided that imprecise size limits are implementation-dependent not implementation-defined. Jonathan to update and complete document, and propose errata that result. 25/3: No further progress 11/4: Not discussed 15/4: Still in progress. Jonathan will take what he has so far and reword as an erratum. This can be added to experience document 1 and then merged into the DFDL spec draft. Jonathan will try and do that this week. 29/4: Jonathan still making progress, will send what he has by end of week. 6/5: Latest draft sent by Jonathan. Review for next time. Mike will incorporate into spec. ... 3/6: No progress 10/6: Not complete. Decided that next published specification would not include this. 17/6: No update 25/6: No update
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. 25/6: No further progress
233	**Public comment : Formats with bit order reversed (Mike)** 1/10: http://redmine.ogf.org/boards/15/topics/43. Mike to provide words for potential new property for review. 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). 22/10: Mike wants to have a working implementation before closing on this, so marking the public comment as deferred. 31/10: Deferring for now 25/6: Un-deferring as Daffodil has implemented this now and Mike has created an experience document in Redmine (). All to review Mike's document before next call on 8th July, and respond with comments. IBM will hold a review and send consolidated comments.
248	**Discriminators and potential points of uncertainty (Steve)** 28/1: Steve to write up a proposal to prevent a discriminator from behaving in a non-obvious manner when used with a potential point of uncertainty that turns out not to be an actual point of uncertainty. 5/2: Steve sent an email to check whether choice branches, unordered elements and floating elements should always be actual points of uncertainty, as there are times when there is no uncertainty, eg, last choice branch; all floating elements found. It was decided that they are always actual points of uncertainty. To do otherwise will complicate implementations and result in fragile schemas. Steve will proceed with the proposal on that basis. ... 25/3: No further progress 11/4: Proposal sent to mailing list by Steve. Concern that having a potential PoU that in practice can never be an actual PoU is counter intuitive and we are better off saying that for certain occursCountKinds there is no potential PoU. The behaviour is therefore the same as for scalar elements. Means that occursCountKind 'fixed' and occursCountKind 'implicit' with minOccurs=maxOccurs behave differently wrt to discriminators. Steve will reword the proposal accordingly. ... 29/4: No further progress 6/5: Steve came to reword the proposal to say that for certain occursCountKinds there is no

	potential PoU, but it raised an issue. Steve has resent the original proposal with responses to Tim's questions. It is clear that a discriminator inside an array can not leak outside the array because it is evaluated for each occurrence. But should that be expressed by saying that a) all arrays are potential PoUs and a discriminator can't leak outside a PoU, or b) only some arrays are potential PoUs and a discriminator can't leak outside a PoU or an array. Please can WG members review the email and have a position on the wording. 20/5: Tim has reviewed, back with Steve ... 3/6: No progress 10/6: Not complete. Decided that next published specification would not include this. ... 25/6: No 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.
258	**Consider allowing more flexible escapeCharacter schemes (Mike)** 6/5: Motivated by example of an escape character which is active when in front of an in-scope delimiter, but not when in front of another character. 20/5: Can't model Mike's example with current facilities, but Mike's example is a generalisation of a particular MITRE example. Do we really need this? Jonathan to follow up. 3/6: ~~**Closed**~~**.** Jonathan has provided the background to the MITRE example which was really about initiators and terminators. The generalised use case is perhaps speculative, so it was agreed not to change the DFDL spec to handle this unless a concrete use case emerges. 17/6: **Re-opened.** vCard 3.0 (http://tools.ietf.org/html/rfc2426) is an example of a format that exhibits the need for this. Need a proposal to handle this case, and which fits in with the existing extraEscapedCharacters and escapeEscapeCharacter property. Noted that using lengthKind 'pattern' is sometimes a way of working round this kind of thing. 25/6: No progress
260	**Positional and non-positional sequences (All)** 10/6: Spec defines the above but also allows different occursCountKinds within the same sequence which may have different (implied) separatorSuppressionPolicy, which results in a sequence which is a mixture of both. Should this be allowed? If so what are the rules? Can certain combinations be disallowed? 17/6: IBM have discussed internally and will submit a proposal. 25/6: Proposal sent to WG. Initial reaction is that the intended semantic for Positional sequence is option a) - an observer of the raw data can identify an occurrence of an element in the sequence solely by counting separators.
261	**Implied separatorSuppressionPolicy for occursCountKind 'expression' (All)** 10/6: Spec says it is 'never' (positional sequence) but you have to parse to identify the position, so isn't that non-positional? 17/6: Some other issues noted around 'expression' as per email thread. IBM have discussed

	this internally and will submit a proposal. 25/6: Proposal sent to WG. The aim of ensuring the integrity of the data stream is sound, but need to think carefully about the use of dfdl:contentLength() in outputValueCalc to make sure there is no deadlock or inconsistency.
262	**Publish updated specification (All)** 10/6: Start to address Word comments in draft r11. Got to the start of section 12.3 (length properties). 17/6: Extra call held, nearly all comments addressed in draft r12 by Mike. Draft r13 created by Steve for remaining comments plus recent spec errata. Mike aiming to create draft r14 for WG review by next week. 25/6: Mike fixing up the references to ensure hyperlinks present in derived html. Then draft r14 can be issued for WG review.

Closed actions

No	Action

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.
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
251	**Create official error codes (All)** 5/2: 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.

Work items:

No	Item	Owner	Target	Status
045	Resolve public comments and incorporate into spec GFD.207	All	2014-04-30	Pending