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-Nov-14 (Tues)
Meeting Time	16:00 - 17:00

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

OGF DFDL Working Group Call , 25 Nov 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. MIL-STD-2045 schemas updated on GitHub
Mike has simplified these.

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
Mark Frost
Alex Wood
Tim Kimber

Apologies
Andrew Edwards

Minutes

1. Daffodil Open Source Project
Not discussed.

2. MIL-STD-2045 schemas updated on GitHub

Mike has simplified these using the suggestions made a few months ago. Much improved.

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

Meeting closed
17:20 UK

Next regular call
Tues 2nd December @ 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: **275**

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. ... 25/11: No further progress
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. ... 26/8: No progress 2/9: No concrete progress although Steve came across this exact scenario when modelling NACHA Addenda records, and used asserts instead of discriminators to side-step the issue. ... 28/10: No further progress 11/11: Steve has looked into this again and will email findings for next call. 18/11. Proposal sent by Steve. Mike to review for next call. 25/11: Mike's review observed that the success behaviour and the failure behaviour of the discriminator becomes asymmetric. That is, when a discriminator fails, the parser backs out an enclosing actual PoU, but when it succeeds it might not positively resolve that PoU. This is not necessarily a problem, and is better than changing the failure behaviour (which needs to remain the same as that for an assert). Alex pointed out the scenario where two (or more) discriminators existed within a PoU, the behaviour today being that the second would resolve a higher level PoU. It was realised that the proposal to prevent leakage would affect this scenario, and would force rewrite of some existing schemas. It was observed that the proposal was an attempt to introduce 'targeted' PoU resolution by the back door, and that workarounds exist to use OCK 'parsed' or split the array into required and optional sections, or to add a choice that used dfdl:occursIndex. Agreed that Steve would revisit his original motivating EDI example and that a behaviour change would only be pursued if he found a compelling reason to do so.
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. ... 25/11: No further progress
274	**XPaths: Namespaces of element names in the path (Mike)** 23/9: Clarify how path steps are to be interpreted wrt namespace. Mike to look at IBM and MITRE suggestions and how unique/key/keyref paths work. 14/10: Mike has researched this, and will circulate to the WG. 28/10: Email sent by Mike. Review for next call. 11/11: Mike to update his research after responses from Tim. Review for next call. 18/11: Mike still can't find language in XSDL 1.0 spec that definitively states how this works, but the book 'Definitive XML Schema' says *"A child element-type name which must be prefixed if it is in a namespace"*. WG agreed that this rule is the one that should be adopted. It is therefore not possible in a DFDL expression to refer to an element that has a target namespace without the use of a namespace prefix. Mike to figure out how this is incorporated into section 23 for next call. 25/11: With Mike for next call.

Closed actions

No	Action
260	**Positional and non -positional sequences (Steve/Tim)** 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. 8/7: Tim emailed an example which he would like to discuss before WG decides on option a) or b). ... 22/7: No progress 29/7: Discussed background, Tim and Steve to have a position for next call. 4/8: IBM has held meeting, Steve to write up findings 26/8: Revised proposal sent by Steve - please review. Need to decide if included in next published spec. 2/9: Concern from Tim that the proposal still allows 'hybrid' sequences that mix 'positional' and 'non-positional' separator suppression characteristics. The alternative is to prohibit certain combinations of occursCountKind and separatorSuppressionPolicy. If this makes implementation materially easier then it is a possibility, but care must be taken to minimise breakage of existing schemas. Tim to make new proposal. Separately, clarification needed on behaviour of maxOccurs '0' and occursCount '{0}' - see **new action 271.** **...** 16/9: With IBM 23/9: IBM held meeting prior to WG call. Discussion ongoing. 14/10: Went though latest email proposal from IBM. Mike to review in depth for next call, and consider spec updates. ... 11/11: Still with Mike 18/11: Mike sent email outlining spec updates. The schema definition error for OCK = 'parsed' and SSP <> 'anyEmpty' was questioned, but WG agreed that was the correct behaviour on the grounds of simplicity. Steve will email the full list of spec updates for review on next call. 25/11: **Closed**. Spec updates emailed by Steve. Steve noted that the OCK 'stopValue' behaviour was not as per the table, and that when unparsing the SSP of the sequence would make a difference. **Erratum raised - tracked by http://redmine.ogf.org/issues/243.**
271	**Use of maxOccurs '0' (All)** 2/9: Legal in XSDL and DFDL. One use case in XSDL is when deriving a complex type by restriction, where it is used to indicate that an element in the base type must not appear. Another use case could be if a schema undergoes a version revision that removes elements; perhaps clearer if the removed element gets maxOccurs '0' rather than omitting it. Steve has seen an instance of its use in a DFDL schema. So if DFDL continues to support it, need to document the behaviour for the various occursCountKinds and what happens to separator. Also behaviour of occursCount expression that returns 0. Proposal needed. ... 23/9: No progress 14/10: Went though latest email proposal from IBM. Agreed with conclusions. Mike to consider spec updates. ... 11/11: Still with Mike 18/11: Mike sent email outlining spec updates. Steve thinks spec needs to say something for OCKs 'fixed', 'implicit' and 'expression' in section 16, and what effect is on separators in section

	14. Mike to re-evaluate spec updates needed for next call. 25/11: **Closed.** Spec updates emailed by Mike. **Erratum raised - tracked by http://redmine.ogf.org/issues/244.**
272	**Catch 22 with emptyValueDelimiterPolicy for fixed length initiated elements (Steve)** 23/9: One solution is to say that a fixed length element can not ever be empty. Needs more thought. 14/10: To be discussed on next call. 28/10: Steve sent out some more thoughts on this. Saying that a 'fixed length' element where length is not zero implies that an empty representation is not possible seems the right thing here. We would define 'fixed length' as 'explicit' (non-expression) or 'implicit' (simple). Noted that 'fixed length' is used in the spec but not defined! 11/11. Review last email for next call. 18/11: Review last email for next call, particularly the implications of saying that a 'fixed length' element can never have an empty representation. 25/11: **Closed.** Agreed that a 'fixed length' element with non-zero length can not have an empty representation, and the ramifications of that. Note there is no effect on nil representation - note that a nilValue of %ES; is possible because with nilKind 'literalValue' the pad character is trimmed before the nil comparison. **Erratum raised - tracked by http://redmine.ogf.org/issues/245**

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: 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