

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	11-Aug-22 (Thurs)
Meeting Time	16:00 - 17:00

Created by Steve Hanson on 09-Mar-11

Last Modified by Steve Hanson on 11-Aug-22

OGF DFDL Working Group Call, 11th August 2022

Agenda

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

1. Implementations 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

Steve Hanson
Mike Beckerle

Apologies

Minutes

1. Implementations Update

Next release to generate output in W3C EXI (binary XML) format.
Also working on generated C parser for small DFDL subset.
Also working on VS Code debugger.

IPR Statement

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

Meeting closed

16:45

8th September 2022 @ 16:00 UK

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.

Actions raised at this meeting

No	Action

No	Action
289	<p>Unparsing: expression refers backwards to outputValueCalc which refers beyond it (Mike).</p> <p>2/8/16: Need to decide if this is allowed and if so if there are any restrictions.</p> <p>13/9: Motivating scenario is where a variable is being set to a length element using dfdl:setVariable, which on unparse is set using dfdl:outputValueCalc. So although the variable is referring backwards to the length element, it is effectively forward referencing so must block. Mike believes this is unavoidable.</p> <p>11/10: Daffodil has implemented this, Mike to provide scenario.</p>

	<p>8/11: Mike couldn't find example, will continue to look</p> <p>10/1/17: Mike has realised that all the examples were reworked to avoid using variables, hence why can't be found.</p> <p>7/2: Daffodil will soon be implementing dfdl:newVariableInstance which will bring this up again.</p> <p>...</p> <p>17/4/18: Waiting for Daffodil to implement dfdl:newVariableInstance</p> <p>15/5: Daffodil team have supplied an example of this from the PCAP schema. Likely to require a flag on newVariableInstance (or maybe variable declaration) to indicate whether needed on parse, unparse or both.</p> <p>...</p> <p>17/10/19: No further progress</p> <p>12/12: Daffodil starting to implement dfdl:newVariableInstance</p> <p>...</p> <p>16/4/20: No further progress</p> <p>30/4: Daffodil has implemented newVariableInstance so now in a position to look at this again.</p> <p>9/7: No update</p> <p>23/7: Daffodil has implemented newVariableInstance so Mike will look at this again</p> <p>3/9: Daffodil team looking at this and related issues surrounding expression evaluation when unparsing.</p> <p>...</p> <p>29/10: No further progress</p> <p>12/11: Mike has a scenario in the PCAP schema that uses several outputValueCalc. To simplify the expressions, he would like to use a variable with a value set by newVariableInstance. The problem is that the rules for variables are different from those for outputValueCalc, specifically variables can not refer forwards (because they are evaluated on both parse and unparse) whereas that is the point of outputValueCalc. To ease this would require a change to the spec. WG agreed this is not necessary for DFDL 1.0 but a candidate for 2.0.</p> <p>10/12: Daffodil to implement as as an experimental feature. Added to table below. GitHub issue https://github.com/OpenGridForum/DFDL/issues/22.</p> <p>7/1/21: Daffodil have started prototype work on this.</p> <p>18/2: Work continuing, but it looks like forward reference variables are needed as anticipated.</p> <p>18/3: Daffodil has implementation of this nearly complete, which Mike will walk through on the next call.</p> <p>15/4: Mike walked through examples of dfdl:newVariableInstance which refer forward and use 'dfdlx:direction=unparseOnly', using the PCAP schema. Will be part of the experience document.</p> <p>8/7: Material assembled for experience document.</p> <p>14/10: Mike to write this up</p> <p>...</p> <p>11/8/2022: Write up still in progress</p>
322	<p>DFDL 2.0 candidate: Handle embedded XML and JSON in a natural way (Mike)</p> <p>8/7/2021: This was something that was included in the first DFDL 1.0 spec drafts but dropped.</p> <p>14/10: Mike has created GitHub issue https://github.com/OpenGridForum/DFDL/issues/27.</p> <p>24/2/2022: Create an experimental feature document</p> <p>14/7: Waiting for prototype implementation</p> <p>11/8: Apache Daffodil implementation in code review and being tested</p>
324	<p>Erratum: Clarify semantic of array element assert containing subexpression ".[1]" (All)</p> <p>8/7/2021: Erratum needed to state that ".[1]" is equivalent to "../a[1]" and</p>

	<p>"../a[dfdl:occursIndex()]"</p> <p>14/10: Mike has created GitHub issue https://github.com/OpenGridForum/DFDL/issues/26.</p> <p>24/2/2022: Mike to look back through emails for more information.</p> <p>...</p> <p>11/8: No further progress</p>
325	<p>Submit OGF DFDL 1.0 to ISO for adoption as an ISO standard (Mike, Steve)</p> <p>11/11: Alan & Wolfgang attended the call. OGF is looking to re-home its successful standards to more relevant organisations. The proposal is that DFDL moves to ISO. OGF has well-established connections with ISO, and Wolfgang is an attendee at working group meetings. There is a specific ISO process that lets them adopt a public, freely available standard whilst maintaining its public, free status. In contrast to ISO-owned standards which are chargeable. The public, free status continues to apply to ongoing revisions. DFDL feels like a good fit for ISO, noted that DFDL schemas exist for ISO 8583 and EDIFACT. There is no need initially to change DFDL to use ISO templates, this can happen later. Noted that this is not a process for dumping old standards onto ISO, it applies to standards with an active community. The end result is that DFDL continues to evolve and be freely available but there are revisions that are effectively 'blessed' by ISO. An example of this in practice is Open Virtualization Format Specification from DMTF. Adoption of a standard by ISO is ultimately determined by national body votes.</p> <p>Next steps: OGF will write to ISO-IEC JTC1 with the proposal. DFDL WG will reach out to the chair of SC38 WG5 which is considered the most likely place DFDL will end up in ISO. The WG5 chair is Fernando Gebara Filho <fgebara@amazon.com>. DFDL WG to consider attending the appropriate ISO meetings.</p> <p>Useful links:</p> <p>ISO Process: https://www.iec.ch/members_experts/refdocs/iec/Consolidated_JTC1_Supplement_2020_publication.pdf</p> <p>DMTF ANSI/ISO adoptions: https://www.dmtf.org/about/register/apresources</p> <p>24/2/2022: OGF are still working on the first step of the process, the application for OGF becoming a PAS submitter. They expect to submit the application letter begin of March. Once OGF will be a recognised PAS submitter there will be an assigned mentor who will guide through the next steps which mostly will have to be taken by the DFDL WG. Wolfgang sent an email with questions that the WG need to answer. Questions discussed and answers sent back to Wolfgang.</p> <p>7/4: Steve to chase OGF for latest position.</p> <p>14/7: From Wolfgang: "The process was a bit slower than foreseen, mostly because of holiday absence both on side the of OGF and JTC1. Regarding the application letter I am waiting for the confirmation from our guide that the last version we have sent him on Monday is well suited for our official application as PAS submitter. If we have his confirmation OGF will officially send the application to JTC1. Thereafter it will subject to a voting of the national bodies which my take additional one or two months." Meantime in parallel Steve will contact the WG5 chair.</p> <p>11/8: ISO have received the application from OGF for it to be allowed to be a PAS submitter. Steve has contacted the WG5 chair who is now following progress.</p>
326	<p>Rules for when an assert or discriminator is used on a prefixLengthType simple type (All)</p> <p>24/2/2022: Mike and Steve believe this is allowed but the meaning of relative path components need to be clarified. IBM DFDL does not support asserts on global simple types so unable to test if it works. Apache Daffodil ignores them. A good way to look at the prefixLengthType is as a quasi-element, so assert feels like it should be allowed. If so need to decide if the expression can use absolute paths, relative paths other than '.' and variables.</p>

	<p>7/4: Struggling to think of a use case where the expression is other than '.', meaning the quasi-element. This is a very restricted use, and suitable for an erratum on DFDL 1.0. Mike to write up.</p> <p>Aside: Steve wondered if validation could be used here instead. IBM DFDL has rich enough validation options that effectively make a validation failure behave like a parse error. Apache Daffodil does not though.</p> <p>14/7: Mike has created a tracker https://github.com/OpenGridForum/DFDL/issues/29. Following from Steve's suggestion, can you use an assert with <code>dfdl:checkConstraints('.')</code> be applied to the type?</p> <p>11/8: Mike has created a PR for the issue. Steve to review.</p>
327	<p>Do we need a default <code>dfdl:choiceDispatchKey</code> branch? (Steve)</p> <p>14/7/22: To handle a nested choice use case where there is a need to distinguish a bad choice branch key from a good choice branch key that then fails to parse its branch. Steve to reproduce the use case to see if IBM DFDL has a similar problem.</p> <p>11/8: IBM DFDL exhibits the same problem. Issue #34 raised. Discussed options for the syntax to indicate that a branch is the default, and agreed that <code>dfdl:choiceBranchKey=""</code> (empty string) works best. Could be an erratum on 1.0 but probably better as 2.0 candidate.</p>
329	<p>DFDL 2.0 candidate: Additional binary number reps (Mike)</p> <p>14/7/22: Apache Daffodil users are encountering a bunch of unsupported binary number formats. Apache Daffodil would like add support for these into DFDL 1.0 as an experimental feature ahead of DFDL 2.0. To do so requires action 328 to extend the experimental syntax to permit this.</p> <p>11/8: Apache Daffodil has also observed that the same is required for binary calendars, eg, unsigned binary seconds.</p>

Closed actions

No	Action
328	<p>Extend the experimental feature syntax to allow extra enums on existing properties (Mike)</p> <p>14/7/22: The experimental feature document does not cater for additional enums to existing properties. Action 329 is an example of where this would be useful.</p> <p>11/8: Closed. Mike has updated the document, Steve has reviewed.</p>

Deferred actions

No	Action
250	<p>Standardise on a single tdml format for DFDL tests (All)</p> <p>5/2/14: 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.</p> <p>...</p> <p>18/2: No further progress</p> <p>11/3: Mike and Steve discussing the best way to share and cooperate on tdml format.</p> <p>25/3: Discussed the creation of an OGF document that will own and define a standardised tdml format.</p> <p>11/4: Proposal is for the OGF document to define a tdml format without Tresys or IBM copyright statement.</p> <p>15/4: Draft document on Redmine</p> <p>...</p> <p>6/5: No further progress</p>

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

...

3/11: No progress

5/1/16: No progress. Needs more interoperability between implementations to be really useful.

...

25/7/17: No further progress

3/10: No further progress although forthcoming work to add packed/zoned numbers may force <https://redmine.org/issues/339> progress

...

11/12: Expected to look at this in the next month or so

...

4/9/18: No further progress

16/10: Mike has started work on a TDML runner that can drive a pluggable DFDL implementation, in support of interoperability testing, including IBM DFDL.

1/11: Pluggable TDML runner working. On Github at <https://github.com/OpenDFDL/ibmDFDLCrossTester>. Schema resolution for IBM DFDL achieved using its schema resolver feature and pointing it at Daffodil's resolver. IBM DFDL sample uses mark() on its input stream but IBM believes this is not necessary.

15/11: IBM DFDL and Daffodil have dependencies on different releases of ICU. Forcing changes to the TDML runner to isolate the implementations under test.

29/11: Good progress on the TDML runner, see email from Mike. The ibmCrossTestRig is not part of Daffodil (because it links against IBM DFDL), but is open source Apache License v2, and is currently in review at <https://github.com/OpenDFDL/ibmDFDLCrossTester/pull/1>. Steve needs to talk to IBM legal to check this is ok as it currently modifies IBM DFDL sample code.

10/1/19: Daffodil have removed the modified IBM DFDL sample code. Steve to ask whether IBM can donate tests from the existing IBM DFDL test suite.

7/2: If permitted, the tests from IBM can be used to see how the IBM and Daffodil tdmls have diverged.

5/4: IBM are permitted to send Daffodil some example tdmls. Steve to send some to Mike for next call.

...

31/5: No progress

27/6: Steve has sent example TDML files to Mike, under the existing IBM Grant of Copyright agreement with Apache.

11/7: Daffodil to adapt their test runner so that it works with the IBM TDML file variant. Action can be closed when this is done.

...

	<p>29/8: No further progress</p> <p>17/10: Noted that this action had its number accidentally swapped with 242 in October 2016. Renumbered back to 250.</p> <p>...</p> <p>23/7/20: No further progress</p> <p>3/9: Nothing likely to happen for this action in medium term, so moving to deferred actions</p>
316	<p>Proposed new experimental feature 'User defined functions' (Mike)</p> <p>9/1/20: Version 2.5.0. of Daffodil will include this, needs writing up. Does XPath 2.0 provide for this?</p> <p>16/4: Nothing in core XPath 2.0 that allows user-defined functions, that is left to XSL etc. A write-up is here: https://cwiki.apache.org/confluence/display/DAFFODIL/Proposal%3A+Feature+to+SUPPORT+User+Defined+Functions. Mike has an example where values need to be normalised for later comparison. Bradd has an example where floating point precision needs adjustment. Steve is concerned that DFDL is straying too far into the transformation layer, maybe "user-defined-functions-for-DFDL" should be a separate spec?</p> <p>...</p> <p>3/9: No further progress.</p> <p>17/9: Nothing likely to happen for this action in medium term, so moving to deferred actions.</p>

Work items:

No	Item	Owner	Target	Status

Experimental Features:

No	Title	Action	GitHub Tracker	Implementor	Experience Document	Release
1	Experimental feature mechanism	301	N/A	Apache Daffodil	https://github.com/OpenGridForum/DFDL/blob/master/docs/current/gwdi-experimental-features-conventions.docx	1.0
2	Elements of type 'xs:hexBinary' with lengthUnits 'bits'	292	https://github.com/OpenGridForum/DFDL/issues/12	Apache Daffodil		2.0 candidate
3	Variable Path Step in DFDL expression	287	https://github.com/OpenGridForum/DFDL/issues/13	ESA DFDL4S		2.0 candidate
4	BLOBs as URIs in infoset	312	https://github.com/OpenGridForum/DFDL/issues/15	Apache Daffodil		2.0 candidate
5	Data streaming layers	304	https://github.com/OpenGridForum/DFDL/issues/11	Apache Daffodil		2.0 candidate
6	Interoperability	307	N/A	Apache Daffodil	https://github.com/OpenGridForum/DFDL/blob/master/docs/current/gwde-dfdl-experience-6-v0.1-interoperability.docx	1.0
7	Empty element	306	N/A	Apache	https://github.com/OpenGridForum/DFDL/blob/master/docs/current/gwde-dfdl-experience-6-v0.1-interoperability.docx	1.0

	parse policy			Daffodil	GridForum/DFDL/blob/master/docs/current/gwde-dfdl-experience-7-empty-ElementParsePolicy.docx	
8	Integer enums as strings in infoset	294	https://github.com/OpenGridForum/DFDL/issues/14	Apache Daffodil		2.0 candidate
10	Indirection using pointers	309	https://github.com/OpenGridForum/DFDL/issues/17	z/TPF DFDL	https://github.com/OpenGridForum/DFDL/blob/master/docs/current/gwde-dfdl-experience-8-experimental-indirection.docx	2.0 candidate
11	User-defined functions	316	https://github.com/OpenGridForum/DFDL/issues/21	Apache Daffodil		2.0 candidate
12	Unparser forward reference variables	289	https://github.com/OpenGridForum/DFDL/issues/22	Apache Daffodil		2.0 candidate
13	Handle embedded XML & JSON in natural way	322	#27	Apache Daffodil		2.0 candidate